ELSEVIER

Soil & Tillage Research 58 (2001) 219–231

Soil & Tillage Research

www.elsevier.com/locate/still

The use of a water budget model and yield maps to characterize water availability in a landscape

Dennis Timlin[a,*], Yakov Pachepsky[b], Charles Walthall[a], Sara Loechel[c]

[a]USDA-ARS Remote Sensing and Modeling Laboratory, Bldg 007, Rm 116, 10300 Baltimore Ave, Beltsville, MD 20705, USA
[b]USDA-ARS Hydrology Laboratory, Bldg 007, Rm 104, 10300 Baltimore Ave, Beltsville, MD 20705, USA
[c]University of Maryland, College Park, MD 20742, USA

Abstract

Crop yield maps may contain substantial corollary information regarding the distribution of yield related soil properties across a landscape. One of these properties is water holding capacity (WHC). Since WHC is an important parameter for crop models and is also critical for crop yield, our objective was to determine if WHC could be estimated by matching simulated yield with yield map data. We collected soil cores for water retention measurements and recorded plant phenological stages from 60 plots on four transects over two growing seasons (1997 and 1998). Soil cores were also sampled on 40 other locations set out on a grid pattern. We utilized a simple water budget model that uses the relative transpiration ratio to calculate relative yield from available water in the soil profile. Rainfall, potential evapotranspiration and soil water holding capacity are input. An optimization program varies the WHC to produce a grain yield similar to the one from the yield map at a particular location. This analysis was carried out over several scales by averaging yields over $55\,m \times 71\,m$, $27\,m \times 35\,m$, and $11\,m \times 14\,m$ areas. Yield data from 2 years were used. Yields from the transects in both years were significantly related to measured WHC in the surface 0–10 cm of soil. The calculated stress indices from the water budget model and estimated available WHC calculated for the 1997 data were similar to those calculated for the 1998 data where data were aggregated in 27 by 35 m or larger blocks. The contour map of estimated WHC was similar to the map of measured WHC for some features though there were also some differences. Use of multiple years of yield data are required to give stable results for estimated water holding capacities.. This information could be used in a farm management plan by allowing a producer to classify a field into areas that are buffered against drought and areas more susceptible to drought. Published by Elsevier Science B.V.

Keywords: Available water; Crop yield; Model; Spatial variability; Precision agriculture; Genetic algorithm

1. Introduction

Spatial information on soil properties and their relationships in a landscape are needed by producers to take advantage of site specific management. Available soil maps, while a source of helpful data, may not have information useful for precision agriculture at scales necessary to successfully implement that technology (Sadler et al., 1998). Crop yield maps, however, may contain a wealth of corollary information about the spatial variability of soil properties that affect yield. Yield maps are also collected at small scales ($\sim$6 m^2) making them an attractive addition to soil maps.

Techniques are needed to interpret yield maps in terms of soil variability and to develop site-specific management practices based on that variability (Mulla, 1991). A typical precision agriculture goal

* Corresponding author. Tel.: +1-301-504-6255; fax: +1-301-504-5823.
E-mail address: dtimlin@asrr.arsusda.gov (D. Timlin).

0167-1987/01/$ – see front matter Published by Elsevier Science B.V.
PII: S0167-1987(00)00170-7

is to use a yield map to divide a field into spatially homogenous sections that can be managed uniformly. This goal may be difficult to realize because there often appears to be a lack of consistency in the patterns of yield variability from year to year (Colvin et al., 1997). To find some basis of consistency, Stafford et al. (1998) applied a fuzzy clustering pattern recognition technique to help classify spatial variability patterns in crop yields. They hypothesized that sub-regions could correspond to areas where yield limiting factors also existed. Further sampling and other tools would be needed to identify and quantify the specific relationships at each location. Overall, there are not a wide variety of tools available.

Among the soil properties that affect crop yields, available water is the most important in rain-fed agriculture. Mean water holding capacity by landscape position has been shown to be a good predictor of corn silage yield (Wright et al., 1990). A large component of variation in yield maps may therefore be due to variation in soil properties that affect water availability. It is often difficult to account for water availability from year to year, however, using statistical approaches. This is because of the temporal stress effects due to limiting soil water and interactions among growth, development and yield. As a method to quantify such effects, Paz et al. (1998) used a crop model to account for seasonal weather effects on soybean (*Glycine max* (L.) Merr.) yields. We hypothesize that information on soil water availability can be extracted from a yield map by using a simple water budget-yield model that can integrate the seasonal effects of weather and available soil water on yield.

The overall goal was to map a surrogate indicator of water availability using detailed crop yields recorded by a yield monitor. Our specific objective was to investigate the use of a simple water budget-yield model to back-calculate soil water holding capacity by matching simulated and measured grain yields within a 4 ha field.

2. Materials and methods

This research is part of an ongoing Precision Agriculture project called OPE3 (Optimizing Production inputs for Economic and Environmental Enhancement) being carried out at the USDA/ARS Beltsville Agricultural Research Center in Beltsville, MD.

2.1. Field

Prior to this study, the field had been in alfalfa (*Medico sativa* L.) for 8 years (1989–1997). In the spring of 1997 the field was sprayed with herbicide to kill the alfalfa, and corn (*Zea Mays* L.) was planted into the residue using a six-row no-till planter. The row spacing was 76.2 cm and plant population about 50,000 plants ha^{-1}. Nitrogen was applied (30 days) after planting at the rate of 100 kg ha^{-1}. No irrigation water was applied.

The soil cover of the site is defined as the Cedartown–Galestown–Matawan soil association (NRCS, 1995). Cedartown and Galestown are sandy, siliceous, mesic Psammentic Hapludults, whereas Matawan is fine loamy, semiactive, mesic Aquic Hapludult. The slopes range from 0 to 5%, and the topsoil texture is loamy sand. Table 1 presents some characteristics of soils.

Table 1
Selected properties of soils in the Cedartown–Galestown–Matawan soil association (NRCS, 1995)[a]

Soil attributes	Cedartown	Galestown	Matawan
Drainage	SED	SED	MWD
Slope (%)	0–5	0–5	0–5
Corn yield (kg ha^{-1})	430	370	740
Thickness of horizons (cm)			
A	3–8	3–5	3–8
E	8–20	5–20	8–18
	Thickness of layers (cm)		
	0–30	0–28	0–50
Clay (%)	2–7	4–10	2–10
Percent passing sieve number			
4	95–100	95–100	100–100
10	85–100	75–100	100–100
40	40–90	45–70	50–75
200	5–25	4–20	15–30
Organic matter (kg kg^{-1})	0.5–2	0.5–2	0.5–2
pH	3.6–5.5	3.6–5.5	3.6–5.5
Available water capacity (%)	5–10	6–8	6–9

[a] SED is somewhat excessively drained, MWD is moderately well drained.

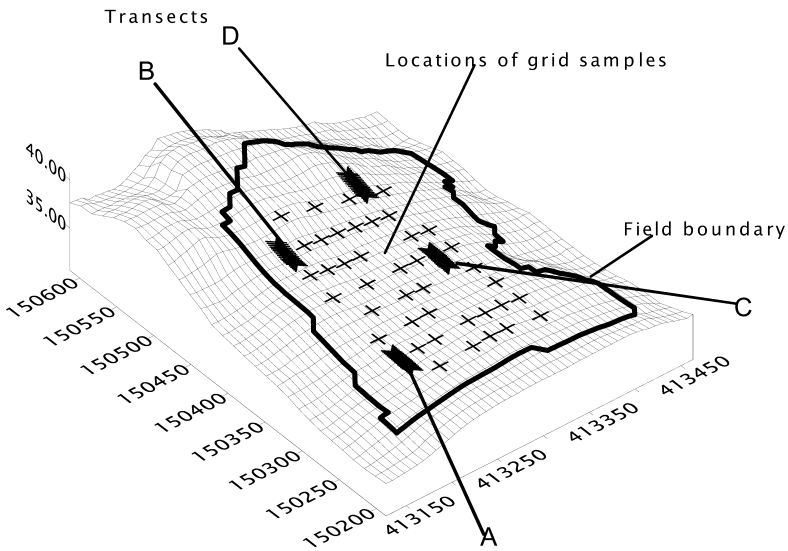

Fig. 1. Elevation map of study site showing locations of transects, grid sampling locations, and field boundary.

An elevation map of the field was obtained by analysis of geo-referenced stereo-pair aerial photographs. Four areas in the landscape where differences in available water could be expected were selected. One transect of 15 plots was laid out in each of the locations. The plot size corresponded to a yield monitor cell ($2\,\mathrm{m} \times 6\,\mathrm{rows}$). An 8×8 grid having an $18\,\mathrm{m} \times 31\,\mathrm{m}$ spacing was also laid out for soil sample sites. A surface map of elevations with the locations of the transects and grid sample locations is shown in Fig. 1.

Crop development (height and vegetative stage) was measured at weekly intervals in the transect plots. Following the 1997 and 1998 growing seasons plot yields were measured manually in all the transect plots. Before the growing season in 1997, duplicate soil cores were collected in each transect plot from 3 to 9 cm in depth. Cores were also taken in every other transect plot (seven per transect) from 24 to 30 cm in depth. Thirty-four duplicate cores (68 total) were collected from 3 to 9 cm depth in the grid. The cores were 6.1 cm high and 5.43 cm in diameter. Moisture retention was measured using a pressure plate apparatus (Klute, 1986) and soil texture measured using the hydrometer method (Gee and Bauder, 1986). The entire field was harvested at the ends of the two growing seasons with a commercial combine equipped with a yield monitor.

2.2. Modeling

A water budget model was used to quantify moisture stress for each yield monitor site. The purpose of the water budget model was to estimate the available soil water content that would give a yield similar to the measured yield if water availability was the only source of yield variability. An optimization program was used to compute the water budget and vary available water content for each measured yield site until the sum of the mean squared differences between predicted and measured yields were minimized.

A simple water budget described by Timlin et al. (1986) was used to calculate a moisture stress index based on available soil water. Corn grain yield predictions are based on the relationship between relative yields and relative transpiration rates (Hanks, 1974):

$$\frac{\text{Actual Yield}\,(Y_\mathrm{a})}{\text{Potential Yield}\,(Y_\mathrm{p})} \sim \frac{\text{Actual Transpiration}\,(T_\mathrm{a})}{\text{Potential Transpiration}\,(T_\mathrm{p})}$$

Here Y_a is the actual yield that results from moisture stress and Y_p is the potential yield where water is not limiting. The water budget calculates values of actual and potential transpiration and weights them for growth stage. The water budget requires weather, soil, and crop data as input. Weather data include daily evapotranspiration, and rainfall. Soil data include available water holding capacity to the bottom of

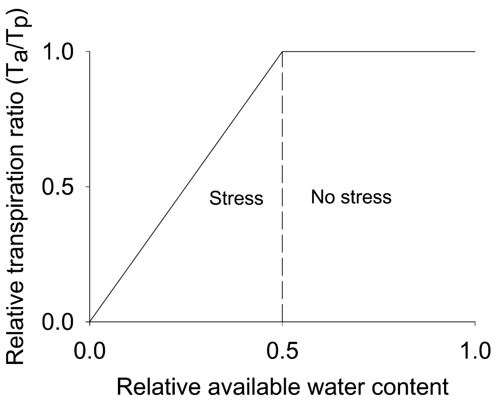

Fig. 2. Schematic of the relationship between the relative transpiration ratio (T_a/T_p) and relative water content. The upper and lower bounds for water content are 10 kPa and 1.5 MPa water contents, respectively.

the root zone and the air-dry water content. Crop data include the number of days after planting that the crop reaches specific growth stages (i.e., vegetative, late vegetative, silking, blister kernel, and maturity) and crop cover coefficients.

The ratio of actual to potential transpiration (transpiration ratio, T_a/T_p) is estimated as a function of the relative available water in the rooting zone (Fig. 2). Relative available water varies between 0 and 1 and is defined as $(w - w_p)/(f_c - w_p)$, where w_p is the water content at -1.5 MPa matric potential and f_c is the upper limit of available water (usually -33 or -10 kPa water content depending on soil texture). Available water is a function of climate and the soil properties that determine water storage capacity. A value of 0.5 for the relative available water content is used as the cutoff point where the relative transpiration ratio becomes <1. As available water content decreases beyond 0.5, the relative transpiration linearly decreases with available water (Fig. 2).

The water budget calculates the current amount of available water in the profile for each day. Water is added to the soil as precipitation and removed as evaporation. The roots grow in the soil profile as a function of time and water uptake from a layer is a function of root density in that layer. Further details can be found in Timlin et al. (1986).

Moisture stress is related to yield through a seasonal moisture stress index (S_s) adapted from Hiler and Clark (1971) and is defined as

$$S_s = \sum_{i=1}^{n} (S_{Di})(W_i)$$

where n is the number of days from planting to harvest, and W_i is a weighting factor that accounts for the sensitivity of grain yield to moisture stress on that day. S_{Di} is the daily stress index for day i as calculated from

$$S_{Di} = 1 - \frac{T_a}{T_p} \tag{1}$$

and T_a/T_p is the relative transpiration ratio.

Hiler and Clark (1971) and Shaw (1974) have used this stress index to calculate corn grain yield using

$$\text{Yield} = Y_p - (A)(S_s) \tag{2}$$

where Y_p is the potential yield when moisture is not limiting and A is the change in corn grain yields in g m^{-2} per unit of seasonal stress (S_s from Eq. (1)). In order to be able to represent results from different sites on a more general basis we will redefine Eq. (2) as a relative yield equation by dividing Eq. (2) by potential yield (Y_p). The result is

$$Y_r = 1 - (A_r)(S_s) \tag{3}$$

where $Y_r = Y/Y_p$, and $A_r = A/Y_p$ is the relative stress index coefficient.

In order to represent the measured yields as relative yields, a value of potential yield is needed. Potential yield, Y_p for each season was estimated by running the water budget using measured soil data from the transect plots with the highest yields (transect A in Fig. 1). The relative yield obtained from the water budget was used with the measured yields to obtain potential yield using the relationships given with Eq. (3). The relative yields for the field data were then obtained by taking the yields measured by the grain yield monitor and dividing by the potential yield. The optimization program minimized:

$$\sum_{i=1}^{n} (Y_r - \hat{Y}_r)^2$$

where Y_r is the measured relative yield and $\hat{Y}_r$ is the estimated relative yield.

2.3. Spatial

The field was divided into sections in order to aggregate data for the optimization method. Three levels of aggregation were used. These three levels were obtained by overlaying grids on a map of the field. The grid spacings used were 55 m (Easting) × 71 m (Northing) (coarse resolution), 27 m × 35 m — (medium resolution), and 11 m × 14 m (fine resolution). Fig. 3 shows a 27 m × 35 m, medium resolution grid overlying the field along with yield monitor locations. All the yields inside a grid cell were averaged and passed to the optimization routine. If there were <5 yield values in a cell the data for that cell were not used in the optimization.

2.4. Optimization method

The optimization program runs the water budget/yield model and varies the WHC of the layers to minimize the differences between simulated and measured yields. The water budget and corn yield simulation model produces a value of relative grain yield for a particular WHC and weather data (precipitation and evapotranspiration). The WHC was varied by changing the amount of water held within the rooting depth. The soil rooting depth was fixed at 90 cm for all locations and was divided into three layers with fixed lower depths (10, 30 and 90 cm) to minimize the number of variables to be fit. Water holding capacities were constrained so that: $\mathrm{WHC}(0-10) > \mathrm{WHC}(10-30) > \mathrm{WHC}(30-90)$.

Fig. 3. Yield monitor locations and layout of 27 m × 35 m grid.

The optimization program we chose for this purpose utilizes a genetic algorithm (GA). Genetic algorithms emulate evolution, and are commonly used as optimization tools in engineering problems. Advantages of GAs include: (1) initial estimates are not important (2) several optimum parameter sets can be found if they exist, and (3) the number of parameters to optimize can be very large (up to hundreds). Genetic algorithms use similar terminology as biological genetics, i.e., chromosomes, genes, alleles, individuals, organisms, and offspring for example. This does not mean to imply any type of physical breeding process or genetic manipulation. Genetic algorithms are only very approximate analogs of biology.

The parameters to be optimized are numbers and can be represented by binary strings, i.e., 0011010. Each binary string representation of a set of parameters to obtain a value of WHC for a grid cell is known as a chromosome. Both floating point and integer numbers can be represented. For example, the water holding capacities of the 0–10, 10–30 and 30–90 cm layers would be considered parameters for the model to calculate a value of relative yield. The binary coding of these three parameters would be available as a string of 0's and 1's in a single chromosome. To calculate a map of WHC for the field we would need a value of WHC for each grid cell in the map. The genetic algorithm would use one chromosome of 0's and 1's to code for a single grid cell. Multiple chromosomes would be used to code for the WHC for the field. These multiple chromosomes would be grouped in an individual or in biological terms, an organism.

Each individual can be assigned a fitness value depending on how well the simulated relative yields agree with the measured relative yields for all the grid cells in the field. To calculate a fitness value, the chromosome is decoded back into floating point numbers that represent the WHCs of the three layers. These WHCs are used in the water budget/yield model to calculate a value of relative yield. This process is iterated over all the grid cells in the field. If the simulated relative yields for all the grid cells were relatively close to the measured ones, the fitness criterion of the individual would be considered to be high. In practice, a WHC for an aggregated grid cell (i.e., 55 m × 71 m) is chosen. The water budget calculates a relative yield for the cell. This process continues for all the grid cells in the field. The fitness value is

$$\sum_{i=1}^{n}[(Y_r - \hat{Y}_r)^2]^{-1}$$

where Y_r is the measured mean relative yield for the grid cell, $\hat{Y}_r$ the estimated relative yield and n is the number of grid cells for the field. The genetic algorithm may have over 100 individuals, each coded to provide a complete set of WHCs for all the grid cells.

Genetic algorithms employ methods for moving from one population of chromosomes (e.g., strings of 1s and 0s, or bits) to a new population by using a kind of "natural selection" together with genetics' inspired operators of crossover and mutation. Each chromosome consists of "genes" (e.g., bits). Optimization criterion becomes a fitness measure of the chromosomes. The selection operator chooses those chromosomes that will be allowed to reproduce, and on average the fitter chromosomes produce more offspring than the less fit ones. Crossover exchanges sub-parts of two chromosomes, roughly mimicking biological recombination between two single-chromosome organisms. Mutation changes the values (0s and 1s) in some genes. Each iteration of selection, crossover, and mutation is called a generation. Usually, hundreds of generations are needed to produce a chromosome with a good fitness. More information on genetic algorithms can be found in Goldberg (1989).

We used a FORTRAN version of a genetic algorithm called GAFORTRAN.[1] This implementation uses tournament selection, in which all chromosomes have equal chances to compete for becoming a parent, but the fitter of any two has a larger probability to become a parent. Elitism is allowed which means that the algorithm is forced to allow to the best individuals to become parents in each generation. The crossover is uniform and the allele exchange occurs in each bit position. To preserve the diversity in the population, the individuals that are similar to many other individuals are punished, and individuals that are different are rewarded. The following parameters were used for this implementation of the genetic algorithm: a population size of 10,300 generations, 0.1 probability of mutation, uniform crossover, two offspring per crossover, and 0.5 probability of a crossover.

3. Results and discussion

Fig. 4 shows cumulative rainfall and evapotranspiration for the 2 years of data. Weather conditions were drier than normal for Beltsville, MD. There was more early season rainfall in 1998 and the drought began later in the season in 1998 so yields were higher in 1998 than in 1997. Earlier planting in 1998 than in 1997 also helped avoid the worst effects of the drought in 1998. Seed emergence was good in both years and the plant populations measured at harvest in the four

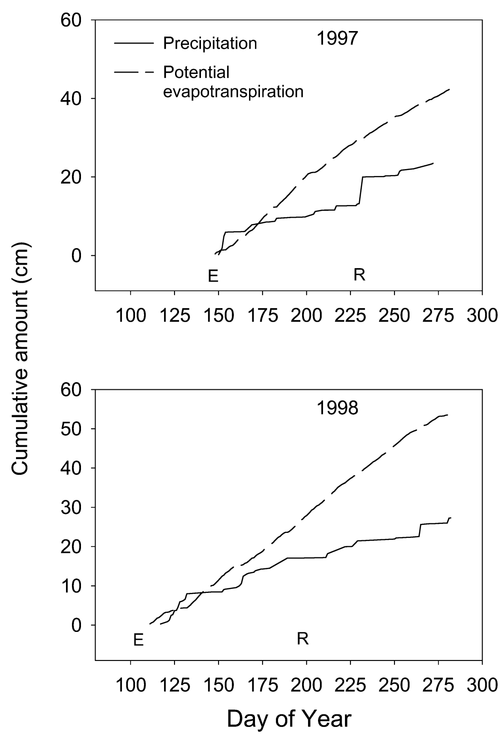

Fig. 4. Cumulative evapotranspiration and cumulative rainfall for the Beltsville site in 1997 and 1998.

[1] David L. Carroll, University of Illinois.

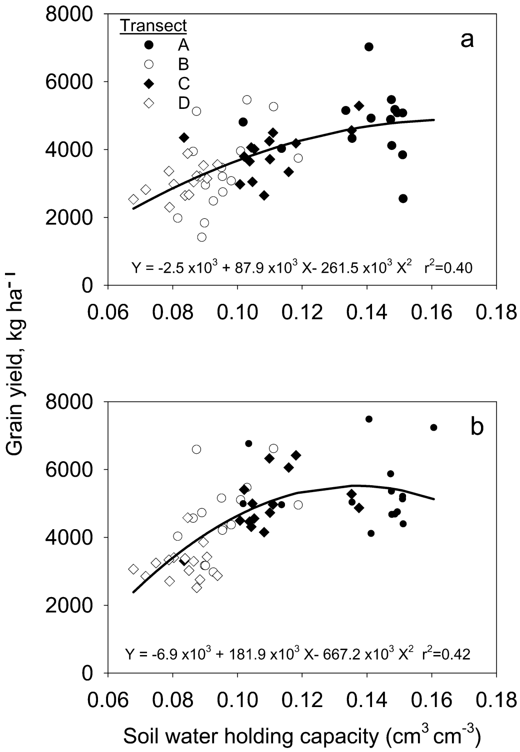

Fig. 5. Relationship between manually harvested transect yield and available water in the upper 10 cm for 1997 (a) and 1998 (b). Open symbols represent transects in the upper (north) part of the field, closed symbols represent plots in the lower (south) part of the field.

transects in 1998 varied from 46,000 to 53,000 plants ha^{-1}. Transect D had the lowest plant population and the only plant population that was significantly less than 50,000 plants ha^{-1} ($^{*}p = 0.05$).

Fig. 5 shows the relationship between available water holding capacity (defined as the difference between 10 and 1500 kPa water contents) in the surface 10 cm and crop yield from the transects for the 1997 and 1998 growing seasons. Crop yields are significantly related to available water capacity for both years but the relationship varies by transect location. Generally, the yields in the upper part of the field (transects B and D in Fig. 1) have lower yields while locations in the lower part of the field (transects A and C) have higher yields. The soils in transects A and C have higher proportions of silt and clay in the upper 10 cm than the soils in plots of transects B and D and hence higher water holding capacities (Table 2). Runoff of rainwater due to soil surface crusting is

Table 2
Measured soil texture components for the soil in the transects (Fig. 1 shows the locations of the transects)

Transect	Sand (kg kg^{-1})	Silt (kg kg^{-1})	Clay (kg kg^{-1})
A	0.74	0.19	0.07
B	0.85	0.10	0.05
C	0.84	0.11	0.05
D	0.90	0.07	0.03

expected to be minimal in this field due to the large amount of surface residue in the no-till conditions.

Polynomial second order regression equations were fit to the data in Fig. 5a and b. The equations and lines are given in the figures. Only the 1998 data show a strong plateau effect where increasing available water (>0.11 $cm^3\ cm^{-3}$) within the upper 10 cm of soil does not contribute to yield increases. The declining yield at high available water in Fig. 5b is not realistic and results from use of a quadratic equation. The quadratic term was not significant in 1997 ($p = 0.24$) so the plateau effect cannot be considered significant. The quadratic term was highly significant in 1998, however ($^{**}p = 0.0031$). The linear contribution of available water capacity to yield in 1998 was almost twice that of the linear contribution in 1997 (Fig. 5a and b).

The relationship between available water in the upper 10 cm and yield shown in Fig. 5 is different for the 2 years. This demonstrates how variations in seasonal weather can impact relationships between soil hydraulic properties and crop yield. Crop simulation models that take into account weather variability can help make sense of such data. The relationships also appear to segregate into groups for the various landscape positions. This suggests that relationships among factors with a cause and effect relationship should be studied with techniques that can identify and characterize local components of the variability. Autoregressive and state-space methods offer this possibility (Long, 1998; Nielsen et al., 1998; Wendroth et al., 1992). This is opposed to methods that characterize global variability such as semivariograms. The relationships in Fig. 5 also show a plateau effect (especially for 1998) that indicates a diminishing return due to increases in available water. This type plateau relationship is seen in irrigation experiments (Pang et al., 1997). The available water content where the relationship begins to plateau would be expected to

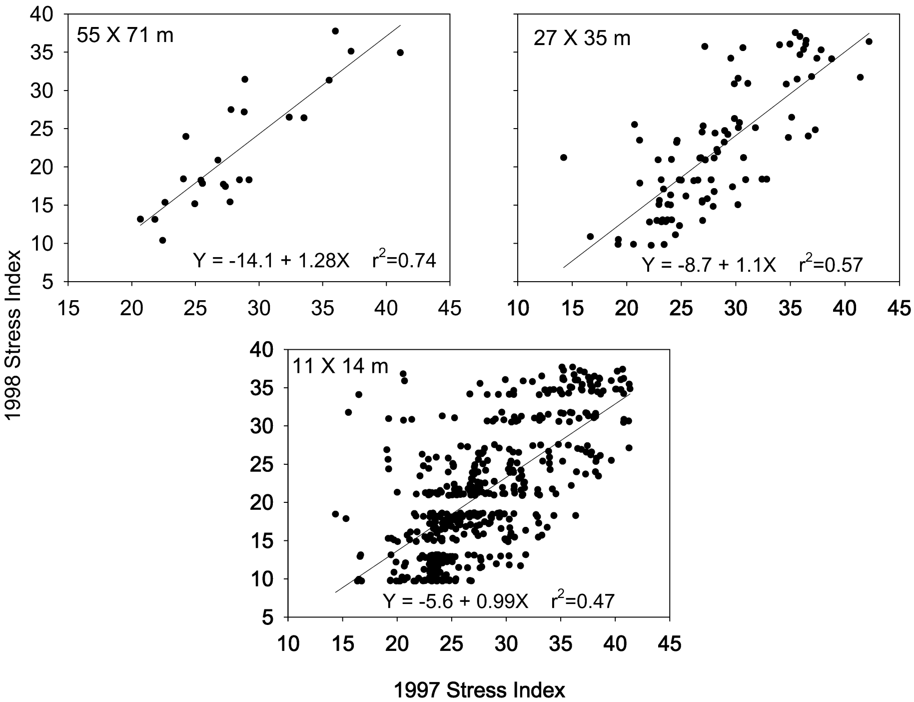

Fig. 6. Relationships between stress indices calculated using 1997 and 1998 yield data for three levels of aggregation (a) 55 m × 71 m, (b) 27 m × 35 m, and (c) 11 m × 14 m.

vary depending on the amount of rainfall and evaporative demand.

The relationships between calculated stress indices for 1997 and 1998 for the three levels of aggregation are shown in Fig. 6. The r^2 and regression equations for the regression of the 1998 stress indices on the 1997 stress indices are given in the figures. The figures suggest that the stress indices are similar across locations, i.e., areas with high stress indices in 1997 also have high indices in 1998. The association between the stress indices for the 2 years increases as more points are added to the aggregation (larger grid cell sizes). Previous research has shown that as the size of the support increases the variability decreases (Zhang et al., 1992). Long (1998) reported that the correlations of wheat yield with normalized reflectance increased as averaging was carried out over larger blocks.

While the stress indices are correlated, the relationship is not 1:1. The stress index at any one location given in Fig. 6 is generally higher for 1997 than for 1998. The field was planted later in 1997 than in 1998. By the time the crop reached the reproductive stages there had been higher PET and less rainfall in 1997 than in 1998 (Fig. 4). The daily precipitation and potential evapotranspiration from planting of the crop to maturity, and soil water holding capacity determine the magnitude of the stress indices. These results suggest that crop yield from a particular area in a field can be related from year to year by using a method that takes the effect of weather variability on yield into account (see also Paz et al., 1998).

The relationships between optimized water holding capacities for 1997 and 1998 for the three levels of aggregation are shown in Fig. 7. As in Fig. 6, the association improves as more data are aggregated (cell size becomes larger). The associations for the low (55 m × 71 m) and medium (27 m × 35 m) levels of aggregation are strongest. The regression line for the low (55 m × 71 m) level of aggregation has a slope and intercept that are not significantly different from 1 ($p > 0.57$) and 0, respectively ($p > 0.27$). The slope and intercept are also not significantly different from 1 ($p > 0.09$) and 0 ($p > 0.06$), respectively, for the

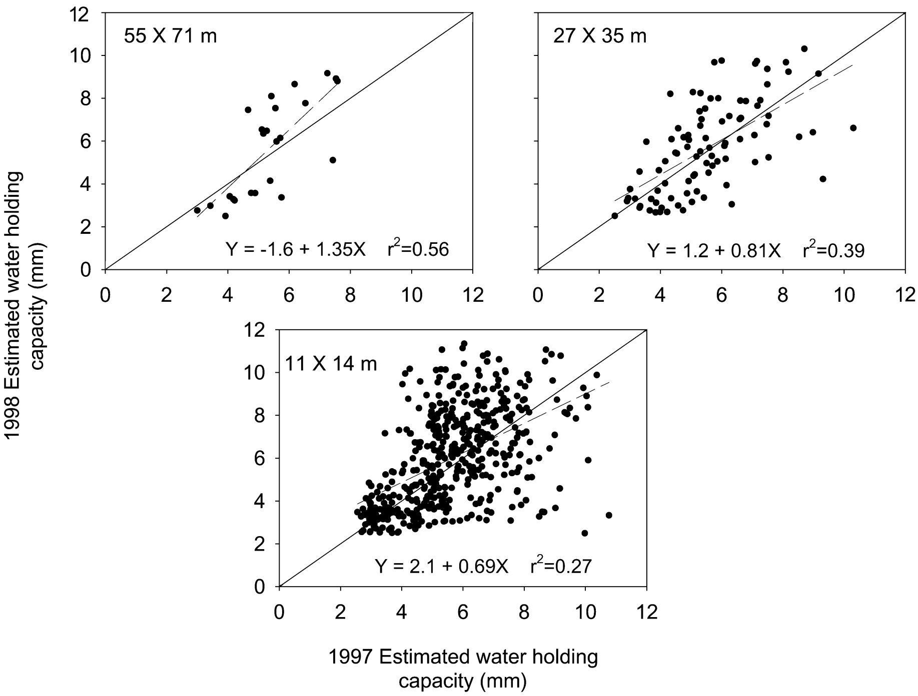

Fig. 7. Relationships between estimated available water using 1997 and 1998 yield data for three levels of aggregation (a) 55 m × 71 m, (b) 27 m × 35 m, and (c) 11 m × 14 m.

27 m × 35 m level of aggregation. The slope and intercept for the 11 m × 14 m level of aggregation were significantly different from 1 and 0, respectively. The regression and 1:1 lines suggest that the relationships do tend to cluster along a 1:1 line though there is considerable scatter. The correlation between the estimated water holding capacities for the 2 years suggest that a reasonably consistent index of water availability can be obtained by using a simple water budget to quantify the effects of soil water availability on crop yield.

The relationships tend to strongly degrade at lower levels of aggregation (smaller cells and larger numbers of cells). The data here show that when the cell size is larger than 27 m × 35 m the correlations using mean data are stronger than they are for the 11 m × 14 m level of aggregation. Based on autocorrelations for the yield data (data not shown) the correlation scale is about 35–40 m. This suggests that the spatial variability might swamp the relationships between the estimated values of available water for the 2 years at the 11 m × 14 m level of aggregation.

Fig. 8 shows contour maps of measured and estimated WHC. Fig. 8a shows a contour map of the measured water holding capacities for the surface 10 cm. Fig. 8b and c shows contour maps for estimated available water from the 1997 and 1998 yield data using the 27 m × 35 m level of aggregation. The estimated water holding capacity is the average value for the 90 cm depth used in the estimation procedure. Fig. 8d shows the mean of the 1997 and 1998 data. The maps are drawn to cover the area of the measured data and so only cover part of the field (see Fig. 1). The patterns for estimated available water for the 1997 and 1998 data are reasonably similar for the 2 years. Regions of high water availability in the lower left hand corner and low water availability in the upper right hand corner are evident. The measured data show similar patterns in the upper right and lower left hand corners of the field. These patterns are related to the soil textures in the surface soil (Table 1). There is more silt in the lower left corner of the field that results in higher available water. The spatial variability of soil water content has been shown to be related to the

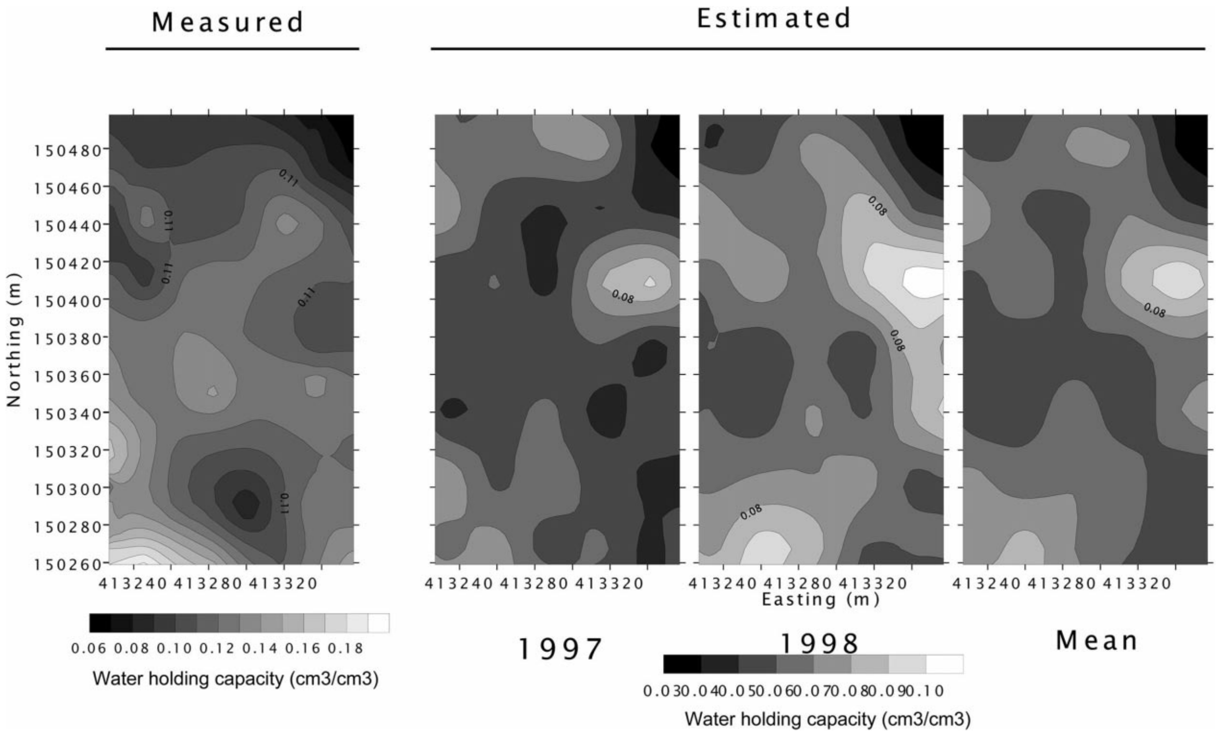

Fig. 8. Contour maps of (a) measured available water holding capacity, (b) estimated available water capacity from 1997 yield data, (c) estimated available water holding capacity from 1998 yield data, and (d) mean estimated available water holding capacity from 1997 and 1998 data. The estimated data are from the 27 m × 35 m level of aggregation.

spatial variability of soil texture (Vauchard et al., 1985).

There is not a strong visual similarity between Fig. 8a (measured) and 8d (mean of estimated values). Patterns, however, are similar. Both maps show that the contours change from southwest to northeast and radiate from the lower left hand corner. There are differences, the map of the mean estimated values shows a region of relatively lower water availability in the middle of the field where the measured data show a region of relatively higher available water. This is the major difference between the measured and estimated maps. We expected some association between measured and estimated WHC since there were significant relationships between WHC in the surface soil and grain yield for the data from the transects.

The measured data are from 0 to 10 cm layer but the estimates are for a simulated 0–90 cm layer. It could be argued that these two may not be directly comparable. Many soil properties at different depths are often correlated with each other (Zhang et al., 1992), however. Surface soil water contents have also been shown to be related to crop yields (Wright et al., 1990; Diaz Zorita et al., 1999). This is due to the higher levels of soil organic matter in the surface soil. Surface soil content is also much easier to measure and hence allows for a larger number of samples to characterize spatial variability. Nielsen et al. (1996) discuss the importance of measuring surface soil water regimes with respect to understanding the spatial and temporal variability of soils.

There are still a number of sources of variability in this method. The available water measurements from the soil cores do not account for soil depth nor for water movement from groundwater. Variations due to soil depth or groundwater would be expected to be localized. The map for the measured data was also interpolated using results from 40 soil cores. Since there were 100 estimated values, the maps may differ because of the different number of samples and the different support size for the yield and soil core measurements.

The effect of diminishing returns will also limit the sensitivity of the optimization method when water availability is above a threshold value. From Fig. 4a and b we could choose a threshold value WHC between 0.10 and 0.12 $cm^3 cm^{-3}$ of water. The optimization method would be sensitive to values of WHC below this value but arbitrarily assign a value of WHC above this point. Furthermore, water availability is unlikely to account for 100% of the variability in yields. Paz et al. (1998) reported that a crop model could explain 69% of the spatial and temporal variability in soybean yields in Iowa.

The ranges of estimated and measured available water content shown in Fig. 8 overlap, although the estimated values are smaller than measured. The estimated available water holding capacity is no larger than about 0.10 $cm^3 cm^{-3}$ and the range is less than in the measured data. This discrepancy most likely because the map shows a mean WHC for 90 cm. The measured values are for the 0–10 cm layer, where WHC are higher than in the sub-surface layers. Mean profile values of WHC would be expected to be lower than values for the surface. This may also be partially attributed to the plateau effect mentioned above. The optimization method is unlikely to choose a value of available water holding capacity greater than this value because the yields are not sensitive to water availability above this level. Another reason for the differences may be partially related to the uncertainty in the definition of available water from the data on soil water retention. The convention is to use capillary pressure of 10 kPa for sandy soils and 33 kPa for all others. A broad generalization has been made when these limits were established. Much smaller values of the ‘‘measured’’ available water content would be obtained if 33 kPa were used as to find the upper limit of the available water.

The area in the right side of the field where estimated WHC is high but the measured value low may represent an area where there is a contribution of water from groundwater. A nearby field has been mapped with ground penetrating radar that has been used to identify the depth distribution of an underlying clay layer. The data suggest that there is lateral movement of groundwater that may be channeled to the surface in some areas of the field (Tim Gish, unpublished data). These kind of hydrologic areas may not be easily identified using moisture retention measurements on soil core samples but still may have a significant impact on yield that may be identified using yield map data.

The level of aggregation impacted the results. As more detail was incorporated into the estimation method, errors had larger effect. Furthermore, when the optimization was carried out the spacial correlations were not taken into account. Doing so would help smooth the relationships. There are a number of methods to do this, though little work has been done on methods that use results from a simulation model. The spatial structure of the water holding capacity is not as easy to obtain as the spatial structure of the crop yields. The spatial variance structure of the crop yields may be used as a surrogate for that of the water holding capacity. Also, the use of elevation information can be used to account for landscape features such as curvature that can affect crop yield (Timlin et al., 1998). Such information could potentially improve this method.

4. Conclusions

The observed scale effect on the correlation between the AWC and stress index in two consecutive years shows two interesting topics to explore. First, point-to-point correlation seems not to be sufficient to express existing dependencies at fine resolutions (e.g., 25×25 scale), statistics taking into account the spatial positions of points should be used. Groovaerts (1999) gives an example of using co-kriging to characterize a spatially variable relationship between rainfall and elevation. Second, the noise manifesting itself at fine scales may preclude using these resolutions in mapping soil properties and delineating management zones. An increase in support mitigates noise and diminishes the overall variability, so that more stable relationships can be established.

The estimated available water holding capacities, however, should still have value despite these issues. This method could provide a good indicator of available water because factors such as sub-surface hydrologic conditions may be accounted for by using several years of yield maps. The estimation method assumes there is no runoff. If runoff does occur its effect will be seen in the optimized water availability since the run-on location may have higher yields than might be possible with no additional water. These optimized available water capacities also provide information on the stability of yield given different yearly weather patterns. While relationships between available water capacity and crop yield can and have been characterized with greater precision on a smaller scale, it is difficult to extrapolate the results to larger scales such as field. The use of crop yield maps in conjunction with a crop simulation model can help identify locations in the landscape where other sources of variability (such as curvature or groundwater) may be important. This will help extrapolate to larger scales. Also, leaf area index from remotely sensed data can be used to better 'fix' crop cover as a function of position and water stress. This kind of information will be included in future research.

Soil map data do contain useful information on water availability but are often not available at resolutions small enough to explain fine-scale variations in yields. Sadler et al. (1998) noted that 1:1200 soil map scales were too weak to be useful for delineating management zones. County soil map scales are 1:20,000. The Beltsville Agricultural Research Center soil map (NRCS, 1995) shows the entire field to be one soil association (Cedartown–Galestown–Matawan). This information cannot be used to explain the yield map variations except at a very gross scale.

In practice, this method would be used with more than 2 years of yield data. The water holding capacity maps would be averaged over a number of years to produce a more stable representation of water availability. Spatial patterns of soil water content have been shown to be temporally persistent (Rolston et al., 1991). It follows that crop yields which depend on stored soil water should show a similar persistence in time. Our work here has shown that a crop simulation model can be used to quantify the effects of soil water on yield and elucidate the temporal persistence of crop yields as they depend on soil water. This information could be used in a farm management plan by allowing a producer to classify a field into areas that are buffered against drought and areas more susceptible to drought. A producer could then target higher levels of inputs to the areas that are more drought resistant and reduce inputs to areas where drought stress is likely. Since the nutrients will be placed where uptake is more likely, the probability of leaching to groundwater can be reduced.

Acknowledgements

The authors wish to acknowledge the assistance of Mr. Dan Shirley of the Beltsville Research Centers Farm Operations Branch with yield data collection and field operations, Mr. Jackson Fisher for data collection, and Mr. Wayne Dulaney for the elevation data. The authors also wish to thank Dr. John Sadler for many helpful suggestions while preparing the manuscript.

References

Colvin, T.S., Jaynes, D.B., Karlen, D.L., Laird, D.A., Ambuel, J.R., 1997. Yield variability within a central Iowa field. Trans. ASAE 40 (4), 883–889.

Diaz Zorita, M., Buschiazzo, D.E., Peinemann, N., 1999. Soil organic matter and wheat productivity in the semiarid Argentine Pampas. Agron. J. 91, 276–279.

Gee, G.W., Bauder, J.W., 1986. Particle-size analysis. In: Black, C.A., et al. (Eds.), Methods of Soil Analysis, Part 1. Agron. Monogr. 9. ASA and SSSA, Madison, WI, pp. 383–411.

Goldberg, D.E., 1989. Genetic Algorithms in Search, Optimization, and Machine Learning. Addison-Wesley, Reading, MA, 412 pp.

Groovaerts, P., 1999. Using elevation to aid the geostatistical mapping of rainfall erosivity. Catena 34 (3–4), 227–242.

Hanks, R.J., 1974. Model for predicting plant yield as influenced by water use. Agron. J. 66, 660–665.

Hiler, E.A., Clark, R.N., 1971. Stress day index to characterize effects of water stress on crop yields. Trans. ASAE 14, 757–761.

Klute, A., 1986. Water retention: laboratory methods. In: Klute, A. (Ed.), Methods of Soil Analysis, Part 1, 2nd Edition. Agron. Monogr. 9. ASA and SSSA, Madison, WI, pp. 635–662.

Long, D.S., 1998. Spatial autoregression modeling of site-specific wheat. Geoderma 85, 181–197.

Mulla, D.J., 1991. Using geostatistics and GIS to manage spatial patterns in soil fertility. Automated agriculture for the 21st century. In: Proceedings of the 1991 Symposium, Chicago, IL, December 16–17, 1991. ASAE Publication, pp. 336–345.

NRCS, 1995. Beltsville Agricultural Research Centre. Special Soil Report. USDA Office of Communications. Washington, D.C.

Nielsen, D.R., Kutilek, M., Parlange, M.B., 1996. Surface soil water content regimes: opportunities in soil science. J. Hydrol. 184 (1/2), 35–55.

Nielsen, D.R., Wendroth, O., Pierce, F.J., 1998. Emerging concepts for solving the enigma of precision farming research. In: Robert, P.C., Rust, R.H., Larson, W.E. (Eds.), Precision Agriculture. Proceedings of the Fourth International Conference, St. Paul, MN, July 19–22, 1998, pp. 303–317.

Pang, X.P., Letey, J., Wu, L., 1997. Irrigation quantity and uniformity and nitrogen application effects on crop yield and nitrogen leaching. Soil Sci. Soc. Am. J. 61, 257–261.

Paz, J.O., Batchelor, W.D., Colvin, T.S., Logsdon, S.D., Kaspar, T.C., Karlen, D.L., 1998. Analysis of water stress effects causing spatial yield variability in soybeans. Trans. ASAE 41 (5), 1527–1534.

Rolston, D.E., Biggar, J.W., Nightingale, H.I., 1991. Temporal persistence of spatial soil–water patterns under trickle irrigation. Irrig. Sci. 12 (4), 181–186.

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

Shaw, R.H., 1974. A weighted moisture-stress index for corn in Iowa. Iowa State J. Res. 49, 101–114.

Stafford, J.V., Lark, R.M., Bolam, H.C., 1998. Using yield maps to regionalize fields into potential management units. In: Robert, P.C., Rust, R.H., Larson, W.E. (Eds.), Precision Agriculture. Proceedings of the Fourth International Conference, St. Paul, MN, July 19–22, 1998, pp. 225–237 and 303–317.

Timlin, D.J., Bryant, R.B., Snyder, V.A., Wagenet, R.J., 1986. Modeling corn grain yield in relation to soil erosion using a water budget approach. Soil Sci. Soc. Am. J. 50, 718–723.

Timlin, D.J., Pachepsky, Ya., Snyder, V.A., Bryant, R.B., 1998. Spatial and temporal variability of corn grain yield on a hill slope. Soil Sci. Soc. Am. J. 62, 764–773.

Vauchard, G., Passerat, A., De Silans, P.B., Vauclin, M., 1985. Temporal stability of spatially measured soil water probability density function. Soil Sci. Soc. Am. J. 49, 822–828.

Wendroth, O., Katul, G.G., Parlange, M.B., Puente, C.E., Nielsen, D.R., 1992. State space approach to spatial variability of crop yield. Soil Soc. Am. J. 56, 801–807.

Wright, R.J., Boyer, D.G., Winant, M.M., Perry, H.D., 1990. The influence of soil factors on yield differences among landscape positions in an Appalachian cornfield. Soil Sci. 149 (5), 375–382.

Zhang, R., Myers, D.E., Warrick, A.W., 1992. Estimation of the spatial distribution of soil chemicals using pseudo-cross-variograms. Soil Soc. Am. J. 56, 1444–1452.