Retrieving Soil Water Contents from Soil Temperature Measurements by Using Linear Regression

Qin XU*[1] and Binbin ZHOU[2]

[1]*NOAA, National Severe Storms Laboratory, Norman, Oklahoma, USA*

[2]*NOAA/NWS/National Centers for Environmental Prediction, Camp Springs, Maryland, USA*

(Received 8 January 2003; revised 18 July 2003)

ABSTRACT

A simple linear regression method is developed to retrieve daily averaged soil water content from diurnal variations of soil temperature measured at three or more depths. The method is applied to Oklahoma Mesonet soil temperature data collected at the depths of 5, 10, and 30 cm during 11–20 June 1995. The retrieved bulk soil water contents are compared with direct measurements for one pair of nearly collocated Mesonet and ARM stations and also compared with the retrievals of a previous method at 14 enhanced Oklahoma Mesonet stations. The results show that the current method gives more persistent retrievals than the previous method. The method is also applied to Oklahoma Mesonet soil temperature data collected at the depths of 5, 25, 60, and 75 cm from the Norman site during 20–30 July 1998 and 1–31 July 2000. The retrieved soil water contents are verified by collocated soil water content measurements with rms differences smaller than the soil water observation error (0.05 m^3 m^{-3}). The retrievals are found to be moderately sensitive to random errors (±0.1 K) in the soil temperature observations and errors in the soil type specifications.

Key words: retrieval, soil water content

1. Introduction

Recently, by using equations of an air-vegetation-soil layer coupled model as constraints, a variational method was developed to compute sensible and latent heat fluxes from conventional observations obtained at meteorological surface stations (Zhou and Xu, 1999, henceforth referred to as ZX99). As an extension of the method of Xu et al. (1999), this method can retrieve the top soil water content and surface skin temperature as by-products. The retrieved water content can be reliable only in the daytime. The air-vegetation-soil layer coupled model used in Xu et al. (1999) and ZX99 was a diagnostic model that did not include the time dimension, so the time tendency information in the soil temperature measurements is not utilized. This leaves some opportunities for further improvements. This is the motivation of this study.

In this paper, a linear regression method is developed to retrieve soil water contents from soil temperature measurements at different depths. The key idea is to utilize the relationship between soil water content and the parameters that control the variations of soil temperature (Hillel, 1980; Koorevaar et al., 1983; Campbell, 1985). The critical parameters are the soil heat capacity and thermal conductivity. These parameters can be estimated by fitting the variations of soil temperature with a simple soil heat transfer equation (Hillel, 1980) in which the soil heat capacity and thermal conductivity are related to soil water content. For a given soil type or soil texture, the soil heat capacity and thermal conductivity are monotonic functions of soil water content, so the soil water content can be determined indirectly from the estimated heat capacity and thermal conductivity. This is the physical basis for the simple method developed in this paper.

2. Model equations and method

2.1 *Model equations*

The heat transfer process in soil can be described by the following equation (Hillel, 1980):

$$C(w)\frac{dT}{dt} = \frac{d}{dz}D(w)\frac{dT}{dz}\,, \qquad (1)$$

*E-mail: Qin.Xu@noaa.gov

where T is the soil temperature, w is the soil water content (measurement by volume in $\mathrm{m^3\ m^{-3}}$), $C(w)$ is the soil heat capacity, and $D(w)$ is the soil thermal conductivity. Both the soil heat capacity and thermal conductivity are functions of soil water content, w, and they both depend on the soil texture. In particular,

$$C(w) = \rho_\mathrm{s} c_\mathrm{s}(1 - k_\mathrm{a}) + \rho_\mathrm{w} c_\mathrm{w} w + \rho_\mathrm{a} c_\mathrm{a}(k_\mathrm{a} - w) \ , \quad (2)$$

where k_a is the soil porosity (in $\mathrm{m^3\ m^{-3}}$), and $\rho_\mathrm{s}, \rho_\mathrm{w}$, and ρ_a are the densities and $c_\mathrm{s}, c_\mathrm{w}$, and c_a are the specific heat capacities for soil, liquid water, and air, respectively. The value of $\rho_\mathrm{s} c_\mathrm{s}$ depends on the soil type (see Table 11–5 of Pielke, 1984), while $\rho_\mathrm{w} c_\mathrm{w} = 4.18 \times 10^6\ \mathrm{J\ m^{-3}\ K^{-1}}$ and $\rho_\mathrm{a} c_\mathrm{a} = 1.2 \times 10^3\ \mathrm{J\ m^{-3}\ K^{-1}}$. Because the air heat capacity is much smaller than the soil and water heat capacity, the term associated with the air heat capacity can be neglected and (2) reduces to

$$C(w) = \rho_\mathrm{s} c_\mathrm{s}(1 - w_\mathrm{s}) + \rho_\mathrm{w} c_\mathrm{w} w \ ,$$

where $w_\mathrm{s}(= k_\mathrm{a})$ is the saturated soil volumetric water content.

According to the empirical formulation derived by McCumber and Pielke (1981) based on the data collected by Al Nakshabandi and Kohnke (1965), soil thermal conductivity (in $\mathrm{W\ m^{-1}\ K^{-1}}$) can be related to soil water content by

$$D(\psi) = \begin{cases} 418 \exp(-\lg|\psi| - 2.7) & \text{for} \quad \lg|\psi| \leqslant 5.1 \\ 0.17 & \text{for} \quad \lg|\psi| > 5.1 \ , \end{cases} \quad (3)$$

where

$$\psi = \psi_\mathrm{s} \left(\frac{w}{w_\mathrm{s}} \right)^{-c} \quad (4)$$

is the soil water potential (in cm) representing the work required to extract water from the soil against capillarity and gravity, ψ_s is the saturated soil water potential, and c is the pore size distribution index which varies as a function of soil (Clapp and Hornberger, 1978; also see Table 11–5 of Pielke, 1984).

Time series of dT/dt, dT/dz, and d^2T/dz^2 can be obtained directly from the soil temperature measurements (at different depths). By least-squares fitting equation (1) to these data, the time-averaged value of $D(w)/C(w)$ can be estimated. With certain simplifications, the method of linear regression can be used for the fitting. The time-averaged soil water content can then be retrieved by using (2) and (3). The method is described in the next subsection.

2.2 *Method*

If the variation of w with the depth can be neglected or if only the bulk (vertically averaged) soil water content needs to be retrieved, then (1) reduces to

$$\frac{dT}{dt} = \frac{D(w)}{C(w)} \frac{d^2T}{dz^2} \ . \quad (5)$$

This reduced form suggests that dT/dt can be linearly related to d^2T/dz^2 for a fixed value of the soil water content w. Such a relationship can be evaluated by fitting the values of dT/dt and d^2T/dz^2 sampled over a time period τ into the following linear regression equation:

$$\frac{dT}{dt} = \beta \frac{d^2T}{dz^2} + \gamma \ , \quad (6)$$

where the regression coefficients β and γ are the slope and intercept of the linear fitting, respectively. The sampled values of dT/dt and d^2T/dz^2 can be computed from the sequential measurements of soil temperature at various depths. The regression slope β can be related to $D(w)$ and $C(w)$ by

$$\beta = D(w)/C(w) \ . \quad (7)$$

The regression intercept γ can be viewed as the bias error of the discretized equation (5) with dT/dt and d^2T/dz^2 computed by the soil temperature measurements.

Once β is computed by the linear regression, the soil water content can be given by the inverse of (7) for the given soil type. To see their relationship, the functional form of $D(w)/C(w)$ is plotted in Fig. 2 for each type of soil. As shown, these functional forms and their inverses are nearly linear for most types of soil, so the soil water content can be uniquely determined from the inverse of (7) with a given β value and soil type. It should be noted that the retrieved soil water content estimates the time-averaged value over the retrieval period τ. Since the time step for the finite difference computation is 15 or 30 min, in order to obtain a large enough number of data samples, τ has to be sufficiently large. However, τ should not exceed 24 hours to resolve the day-to-day variations of w. In this paper, τ is chosen to be 24 hours, so the retrieved soil water content is a daily averaged value.

In general, the method requires three or more depths of soil temperature measurements. As we can see from Fig. 1, the diurnal variation of the soil temperature decreases rapidly with the depth and nearly diminishes below a certain depth. In particular, the temperature at –30 cm is approximately equal to the daily averaged temperature at –5 or –10 cm. This suggests that the deep soil temperature (below the measurement levels) could be roughly estimated by the daily averaged value of the available soil temperature measurements. In this case, the method can be extended to the case in which only two levels of soil temperature measurements are available.

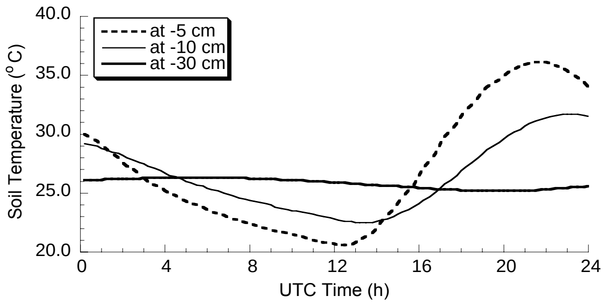


Fig. 1. Daily soil temperature variations measured every 15 min on 11 June 1995 at three different depths at the Oklahoma Mesonet station #38.

2.3 *Equation error and extended method*

The above method uses the simplified equation (5) which is an approximation of (1). To discuss the equation error and extend the above method, it is convenient to rewrite (1) into the following form:

$$P = \frac{D(w)}{C(w)} \frac{d^2T}{dz^2} , \tag{8}$$

where

$$P \equiv \frac{dT}{dt} - Q \text{ and } Q \equiv C^{-1}\frac{dD}{dz}\frac{dT}{dz}$$

is the term neglected in (5). The daily averaged part of Q could just be a part of the total bias error in (5). The total bias error is modeled by the regression intercept γ in (6), but the random part of the equation error is not modeled. The related problem is examined in section 4.3.

Since soil temperatures were measured at four depths at some Mesonet sites (such as the Norman site as described in section 3), w can be estimated for two overlapped layers (an upper layer from –5 to –60 cm and a lower layer from –25 to –75 cm). Hence, the vertical derivative of D in Q can be calculated by the finite difference of D between the two layers. Since C can be computed from the estimated w, and since dT/dz is directly available from the observations, Q can be estimated for each layer at each observation time. Subtracting the estimated Q from dT/dt gives an estimate of P for each layer at each observation time. Clearly, the neglected term Q can be recovered if dT/dt is replaced by P in the linear regression equation (6). This modified equation can be used to fit the above estimated values of P to re-estimate w. The procedure can be repeated iteratively and the solutions are found to converge rapidly (often within 5 steps and never exceeding 50 steps). Although this extended method recovers the neglected term Q, the retrieved soil water contents are improved only slightly in the upper layer but degraded slightly in the lower layer. The mixed result could be due to the fact that the random part of the equation errors is large (see section 4.3) and not modeled by the method. Since the improvement is insignificant, the results of the extended method are omitted in this paper.

2.4 *Observation error and time step*

The soil temperatures measured by thermocouples at the Oklahoma Mesonet stations are accurate only to 0.3 or 0.5 K, but the thermocouples can detect temperature variations of 0.1 K or even smaller. The limited accuracy is mainly due to calibration errors. As calibration errors do not change with time, they are bias errors. All the bias errors do not affect the computed values of dT/dt. They affect the computed values of d^2T/dz^2 and dT/dz but the effects do not change with time and thus can be absorbed by γ in (6) [or in the modified (6) with dT/dt replaced by P]. Consequently, it is only the random part of the temperature observation error that actually affects the retrieved w. In general, an observation error includes two parts: the measurement error and sampling error. The sampling error may be smaller than the measurement error, provided that the spatial variations of soil temperature are small within the volume represented by the sampled value. Since no information is available to quantify the sampling error, the random part of the observation error may be roughly estimated based on the instrument sensitivity, that is, the smallest temperature variation that can be accurately detected by the thermocouple. For data used in this paper, random observation errors could be no more than 0.1 K. Their effects on the retrieved w will be examined in section 5.2.

When the time tendency term dT/dt is computed by the standard finite difference scheme, the accuracy is affected not only by random errors in the soil temperature observations (which could be in the range of ±0.1 K) but also by the time step. The typical value of the soil temperature tendency is about 1 K per hour. The sampling rate of the soil temperature is every 5 minutes at the Oklahoma Mesonet stations, but the data used in this paper were archived every 15 minutes. If the original measurement time step of 5 min is used in the finite difference computation of the time tendency term, then the error in the computed time tendency term could be as large as 0.1 K per 5 min or, equivalently, 1.2 K per hour, which is slightly larger than the typical value (1 K per hour) of the soil temperature tendency. To reliably estimate the time tendency term (with the estimated error not larger than 1 K per hour), the time step for the finite difference computation should be at least 15 min. If the time step is too large, the truncation error caused by the finite difference will become large. In this paper, the time step is chosen to be 15 or 30 min. The Oklahoma Mesonet soil temperature data used in this paper were archived every 15 min for the period of 11–20 June 1995 but every 30 min for the periods of 20–30 July 1998 and 1–31 July 2000.

3. Description of the data

The method will be applied to the same data used in ZX99, so that the retrievals of soil water content can be compared with those obtained from the previous method. The data were collected by the Oklahoma Mesonet during the period of 11–20 June 1995, which were typical summer days without severe storms over the entire mesonet area. Among the 111 Oklahoma Mesonet stations, there were only 50 stations measuring soil temperature at three depths (5, 10, and 30 cm). The soil temperature data measured at these 50 stations can be used to retrieve the soil water contents (in the 5–30 cm layer). The measurement accuracy is about 0.3–0.5 K. As explained in section 2.3, this limited accuracy is mainly due to calibration errors which do not change with time. The random part of the observation error could be in the range of ± 0.1 K. These soil temperature data were archived every 15 min, so there were totally 97 time levels of observations per day. The sampling size for the computed time derivatives is thus 96 and should be sufficiently large for the linear regression. The retrieved soil water content represents the daily averaged value.

The ARM soil water content data used in this paper were collected by the soil water and temperature system (SWATS) installed at the Oklahoma Pawhuska ARM station #12. The soil water contents measured at the ARM station #12 will be used to verify the retrieved soil water contents from the soil temperature measurements at the Oklahoma Mesonet station #38 (about 200 meters away from ARM #12) in section 4.1. The ARM SWATS measures both soil temperature and soil water at the depths of 5, 15, 25, 35, 60, and 85 cm by using the Campbell Scientific 229-L heat dissipation matric potential sensors (Reece, 1996) arrayed vertically into the soil. The soil temperatures were measured accurately to within 0.1 K. To assess the quality and accuracy of SWATS soil water measurements, Schneider et al. (2003) compared the SWATS measurements with those sampled by the benchmark oven-drying method. The assessed accuracy was about 0.05 $m^3\ m^{-3}$ in terms of rms error. Since 132 benchmark measurements were sampled within 5 m of the SWATS instruments for the comparison, their assessed rms error (0.05 $m^3\ m^{-3}$) included both the measurement error and sampling error. Thus, the total observation (measurement plus sampling) error could be very close to the value (0.05 $m^3\ m^{-3}$) assessed by Schneider et al. (2003).

In recent years, the Oklahoma Mesonet has also installed the Campbell Scientific 229-L heat dissipation matric potential sensors at 60 of the 114 stations to measure soil water content profiles at the depths of 5, 25, 60, and 75 cm. Soil temperature and water content data used for the tests of the method in sections 4b–c were collected every 30 min at the four depths from the Norman site for the periods of 20–30 July 1998 and 1–31 July 2000. Soil texture data are also available for the above four depths. The data collected during 1–31 July 2000 will also be used in section 5 to study the sensitivities of the retrievals to random errors in the soil temperatures and to variations in soil types. Since these soil temperature data were archived every 30 min, there were totally 49 time levels of observations per day. The sampling size for the computed time derivatives is thus 48 and should still be sufficiently large for the linear regression.

Basara and Crawford (2000) reported an installation flaw that caused rainwater to flow down the instrument wire and moisten the deep-layer sensors (–60 and –75 cm) at the Norman site. This led to significant errors in deep-layer soil water content measurements as revealed by significant inconsistencies between near-surface (–5 and –25 cm) and deep-layer (–60 and –75 cm) measurements of soil water content during an extended period (about 10 days) after the rainfall of 18 July 1997. Similar inconsistencies are seen from the measurements during the period of 8–18 July after the rainfall on 8 July 1998 at the Norman site. The inconsistencies diminished after 19 July, so the data used in this paper for the dry periods of 20–30 July 1998 and 1–31 July 2000 should not be affected by the reported problem.

4. Applications to Oklahoma Mesonet data

4.1 *Results for the period of 11–20 June 1995*

As a qualitative check, the range of β estimated from the regression of (6) is compared with the range of D/C in Fig. 2. The results (not shown) indicate that all the estimated values of β are well within the valid range of D/C in Fig. 2. When β is estimated from the soil temperature data, the estimated value is independent of soil texture. Soil texture information is required only when the soil water content is further derived from the estimated β.

For a quantitative verification, nearly collocated observation of soil water content is available only at

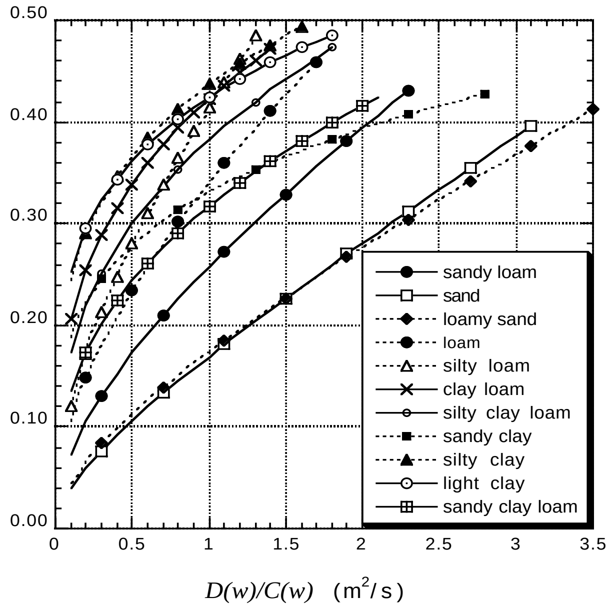


Fig. 2. $D(w)/C(w)$ for different types of soil.

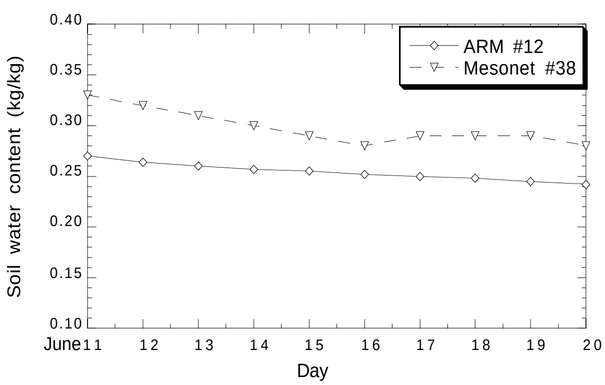


Fig. 3. Comparison between the retrieved soil water content at Mesonet station #38 and the direct measurement at ARM station #12.

one Oklahoma ARM station (ARM #12) which is about 200 meters away from the Oklahoma Mesonet station #38. The elevations of the ARM #12 and Mesonet #38 stations are 331 and 330 m, respectively, and the terrain is quite flat. The surface type can be characterized as native prairie around these two stations. The soil type is sandy loam from the surface to –30 cm at both stations. The retrieved soil water contents (in $m^3\ m^{-3}$) at this station (Mesonet #38) are compared with the direct measurement (with an accuracy within $\pm 0.05\,m^3\ m^{-3}$) at the ARM station (ARM #12) in Fig. 3. As shown, both the observed and retrieved values decrease slowly with time, and there is a nearly constant gap between these two curves. Note that the retrieved soil water content should represent the value in the 5–30 cm soil layer, while the measurement was made for the 0–5 cm soil layer. The gap between these two curves could be largely due to the difference in their represented depths.

Bulk soil water contents are also retrieved for the remaining 49 Oklahoma Mesonet stations. The retrieved soil water contents (in $m^3\ m^{-3}$) can be compared with the retrievals obtained as by-products from the vegetation-soil layer coupled method of ZX99. The method of ZX99 requires measurements of air temperature at two levels (1.5 and 9.0 m). These measurements were available only at the enhanced stations. There were only 14 enhanced stations among the 50 stations where soil temperature was measured at three depths (5, 10, and 30 cm). At these 14 stations, the retrieved soil water contents from the current method are compared with those retrieved by the method of ZX99. The results can be summarized into four groups in terms of the time mean and variance of the difference between the two retrievals. For the first group (stations #3, 93, 98, and 103), the differences are persistently small (see Fig. 4a for station #3). For the second group (stations #10, 11, 29, and 40), the differences are persistently large with small variances (see Fig. 4b for station #10). For the third group (stations #49, 58, and 95), the mean differences are small but the variances are large (see Fig. 4c station #49). For the fourth group (stations #4, 12, and 56), the mean differences and variances are both large (see Fig. 4d for station #4).

The soil water contents retrieved by the current method are very persistent with only small time variations at all the 14 stations. This feature (slow time variation) is consistent with the fact that there was no precipitation over the entire Oklahoma Mesonet area during these 10 days. The soil water contents retrieved by the method of ZX99, however, exhibit strong time variations as shown in Figs. 4c,d. These time variations do not seem to be realistic, because there was no precipitation at any of these stations. These unrealistic time variations could be related to the following two shortcomings in the method of ZX99. First, the method of ZX99 is reliable only for the daytime (between 10 am and 4 pm) with clear or partially cloudy sky. During the nighttime, the canopy resistance reaches the maximum value (in response to the disappearance of solar radiation) and becomes independent of soil moisture in the air-vegetation-soil layer coupled model [see equations (12)–(13) of Xu et al., 1999], so the method cannot (and does not need to) retrieve the soil water content (as a by-product). Even for the daytime, the accuracy of the retrieval may be deteriorated by widespread cloudiness. Secondly, the

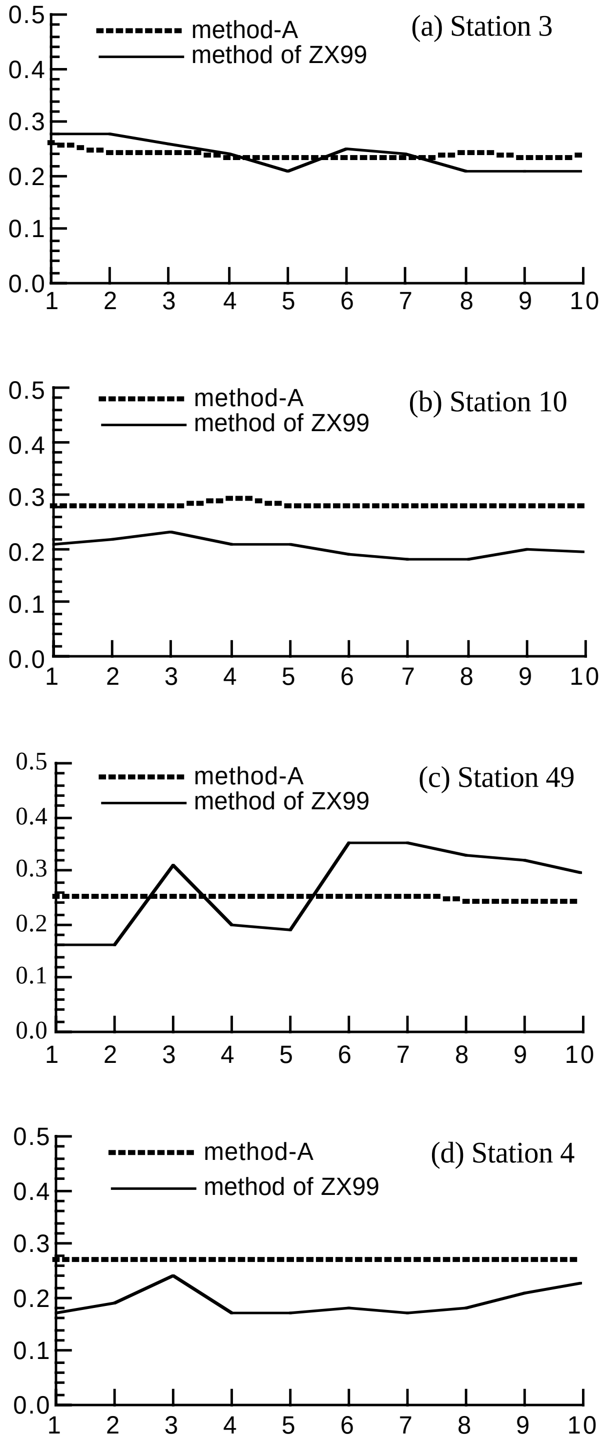


Fig. 4. Soil water contents (in m^3 m^{-3}) retrieved by the current method (dashed) and the method of ZX99 (solid) at Mesonet stations #3 (a), #10 (b), #49 (c) and #4 (d).

air-vegetation-soil layer coupled model contains only diagnostic equations. As the method of ZX99 uses only a single time level of data, the retrieved soil water content is not constrained by any time continuity. These two shortcomings could cause the method of ZX99 to be less reliable than the current method as shown in Figs. 4c and d.

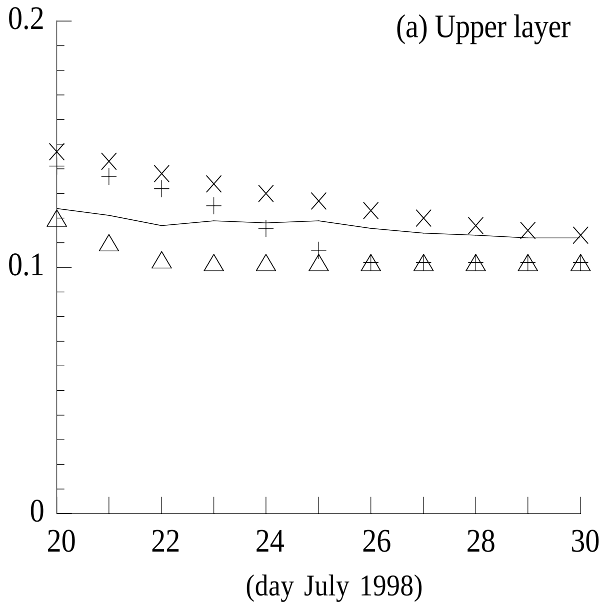
The relatively large mean differences in Figs. 4b and d could be partially attributed to the variation of soil water content with depth. The soil water content retrieved by the method of ZX99 should represent the value in the top soil layer (0–5 cm). The soil water content retrieved by the current method represents the value in the 5–30 cm layer. The differences between the solid and dashed curves in Figs. 4b and d are about 25% of the retrieved soil water contents. The irregular time variations shown by the solid curve in Fig. 4c are also in the range of $\pm 25\%$ of the time mean of the retrieved soil water content. The irregular time variations could be related to the absence of the time dimension in the method of ZX99 (as explained in the introduction section). By using additional soil temperature data, the method developed in this paper produces more persistent retrievals than the method of ZX99.

4.2 *Results for the period of 20–30 July 1998*

As mentioned in section 3, soil temperature and water content data were obtained at the depths of 5, 25, 60, and 75 cm from the Norman site for the period of 20–30 July 1998. Two partially overlapped layers can be considered: an upper layer from –5 to –60 cm, and a lower layer from –25 to –75 cm. Since each layer covers three measurement depths, soil water content can be retrieved for each layer. Based on the soil texture data obtained at the above four depths, the layer-averaged soil type can be specified as silt clay loam for the upper layer and silty clay for the lower layer. The time step used for the finite difference computations in (6) is 30 min, which is the same as the data time interval.

In Fig. 5a, the retrieved upper-layer soil water content (solid) is verified against the direct measurements at the depths of 5, 25, and 60 cm. As shown, the retrieved soil water content is reasonably close to the measurements and is well within the range bounded by the measurements at the three depths. Figure 5b shows that the retrieved lower-layer soil water content (solid) is very close to the direct measurement at –75 cm and is mostly in the narrow range bounded by the measurements at the three depths. These results show that the method works reasonably well for this case.

The weather at the Norman site during the period of 20–30 July 1998 was mostly clear. The soil was continuously dried after the rainfall (0.32 inch) that occurred on 8 July. This drying process is seen from the direct measurements plotted in Figs. 5a–b. The retrieved soil water contents,however,do not exhibit the



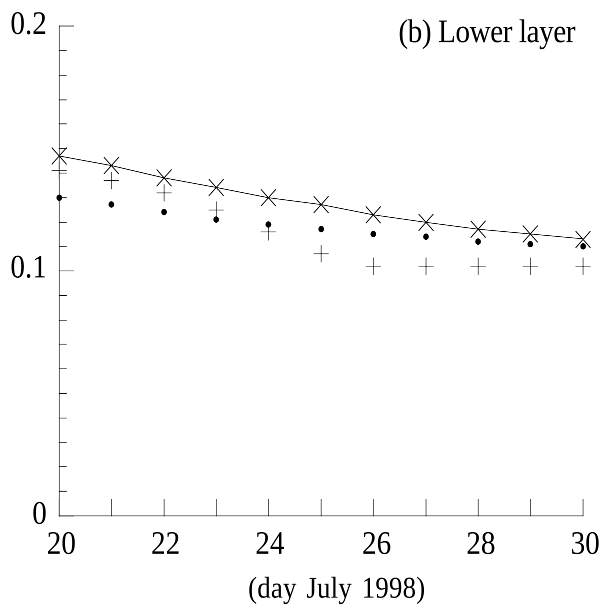


Fig. 5. Retrieved soil water contents (solid) in $m^3\ m^{-3}$ compared with measurements at three depths in the upper layer (a) and lower layer (b) at the Norman site. Measurements are marked by $\triangle$ for 5 cm, $+$ for 25 cm, $\times$ for 60 cm, and $\bullet$ for 75 cm depth.

same drying rates as the measurements. The discrepancy is small but quite persistent through the entire period (11 days).

4.3 *Results for July 2000 and assessment of equation error*

The soil temperature and water content data collected at the depths of 5, 25, 60, and 75 cm from the Norman site for the period of 1–31 July 2000 are used to further test the method. During this period, the observed soil water contents were in a much higher range (between 0.28 and 0.39) than those observed during the period of 20–30 July 1998 (see Fig. 5). This period is thus considered as a wet period. By using the method described in section 2.2, retrievals are obtained for the same two partially overlapped layers as described in section 4.2. The retrieved soil water contents are verified against the observations averaged from three depths in each layer; that is, the depths of 5, 25, and 60 cm for the upper layer, and the depths of 25, 60, and 75 cm for the lower layer. The rms differences between the retrieved soil water contents and averaged observations are listed in column 1 of Table 1. As shown, the rms differences are smaller than the soil water observation error ($0.05\ m^3\ m^{-3}$). These differences are larger than those for the dry period (20–30 July 1998) examined in section 4.2. Thus, the retrievals have relatively large errors for the wet period, but the relative errors (with respect to the observed soil water contents) are about the same for the wet and dry periods. As explained in section 2.3, errors in the retrievals are caused in general by errors in the discrete model equation and input observations of soil temperature and soil type. The related problems are examined below.

The regression intercept γ in (6) represents the bias part of the equation error in the discrete form of (5). The impact of this part can be assessed by comparing the fitted γ with the rms amplitude of dT/dt computed from the observations. Here, dT/dt represents the combined effect of the right-hand side of (5) and the neglected term Q [see (8)]. The averaged values of $|dT/dt|$ and $|\gamma|$ for the entire period of July 2000 are listed in columns 2 and 3 of Table 1. As shown, the averaged $|\gamma|$ is only about 6% of the averaged $|dT/dt|$ in the upper layer, but about 40% in the lower layer. This indicates that the bias error of the discrete form of (5) is small in the upper layer but not small in the lower layer.

As the bias part of the equation error is largely absorbed by the regression intercept γ, the retrieval accuracy may be affected mainly by the random part of the equation error. Limited by the linear regression used in this paper, the random part of the equation error is not modeled. Since dT/dt and d^2T/dz^2 are computed directly from observations in the linear regression equation (6), the equation error is actually mixed with the observation error. Nevertheless, loosely and in the least-square sense, the random part of the equation error can be measured by the rms residual of (6), that is,

$$\sigma = \langle (dT/dt - \beta d^2T/dz^2 - \gamma)^2 \rangle^{1/2} \,, \qquad (9)$$

where $\langle \cdot \rangle$ denotes the assemble average calculated from observations sampled over the retrieval period τ (=one day) based on the assumed ergodicity. The averaged σ for July 2000 is listed for each layer in column 4 of Table 1. As shown, the averaged σ is only about 66% of the averaged rms amplitude of dT/dt in the upper layer, but about 120% in the lower layer. Thus, the random error in the discrete form of (5) is large, especially in the lower layer.

Note that the soil water is retrieved from β by inverting (7), while β is computed as a slope from the linear fitting of (6) to the data cloud of $(d^2T/dz^2, dT/dt)$. The statistical significance of the fitting depends on the sample size and the correlation coefficient between dT/dt and d^2T/dz^2. The correlation coefficient, denoted by R, can be computed from the sampled values of dT/dt and d^2T/dz^2 for each retrieval period. The averaged values of $|R|$ for July 2000 are listed in column 5 of Table 1. Since the sample size is 48, the confidence level will be above 95% if $|R|$ is larger than 0.404. This confidence level (95%) is reached for the upper layer but not for the lower layer. Clearly, the diurnal variations of soil temperature in the lower layer were too weak to allow β to be determined with the required confidence level (95%). Although β can be reliably determined in the upper layer, the soil water content w inverted from β through (7) is not ensured to be accurate unless the soil texture data are accurate, because the inverted w is moderately sensitive to errors in the texture data as shown in the next section.

4.4 *Sensitivity tests*

As explained in section 2.3, the random part of the observation error could be in the range between ± 0.1 K for the conventional soil temperature measurements at the Oklahoma Mesonet stations. The soil temperatures measured by the Campbell Scientific 229-L heat dissipation matric potential sensors at the Norman site were accurate within 0.1 K. Thus, random errors in the range between ± 0.1 K are added to the observed soil temperatures to test the sensitivity of the retrievals to observation errors. The results are listed in Table 2. In comparison with the results in Table 1, the rms differences between the retrieved soil water contents and averaged observations are increased by 19% in the upper layer and by 16% in the lower layer. The retrievals are thus not very sensitive to the observation errors. As random errors are added to the soil temperature measurements, $|dT/dt|$ and σ are expected to increase. This is indeed the case as shown in columns 2, 4, and 5 of Table 2 in comparison with those in Table 1. Table 2 also shows that $|R|$ is decreased in the upper layer, as expected, but increased in the lower layer. The increased $|R|$ in the lower layer is still very small and thus the fitting remains insignificant.

Based on the soil texture data at the Norman site (see section 4.2), the soil types specified for the upper and lower layers are type 7 (silt clay loam) and type 9 (silty clay), respectively (see Fig. 2). These specified soil types may not be as accurate as the original soil texture data. To test the sensitivity of the retrievals to errors in the soil type specifications, the selected soil types are shifted by +1 in their numbers, that is, from type 7 to type 8 (sandy clay) for the upper layer and from type 9 to type 10 (light clay) in the lower layer. Note that the observed soil water contents are in the range between 0.28 and 0.39. Over this range, the shifted soil type 10 is close to the original type 9 in the lower layer, but the shifted soil type 8 is not close to the original type 7 in the upper layer. In response to these changes, the rms differences between the retrieved soil water contents and averaged observations are increased from 0.0336 to 0.0490 in the upper layer and from 0.0275 to 0.0410 in the lower layer.

Table 1. Averaged rms differences between the retrieved soil water contents and averaged observations for July 2000 and averaged values of $|dT/dt|$, $|\gamma|$, σ, and $|R|$

Averaged	rms	$\|dT/dt\|$	$\|\gamma\|$	σ	$\|R\|$
Upper layer	0.0336	0.0637	0.0037	0.0420	0.7557
Lower layer	0.0275	0.0257	0.0104	0.0307	0.0275

Table 2. As in Table 1 but with random errors (between ± 0.1 K) added to the observed soil temperatures

Averaged	rms	$\|dT/dt\|$	$\|\gamma\|$	σ	$\|R\|$
Upper layer	0.0400	0.0915	0.0037	0.0955	0.4798
Lower layer	0.0319	0.0749	0.0079	0.0892	0.0430

Note that the soil water content is obtained by inverting (7) based on the specified soil type for each layer only in the last step after β is computed by the linear regression (see section 2.2). The linear regression of (6) and its derived values of β, γ, σ, and R are independent of the soil type specifications, so all the values listed in columns 2–5 of Table 1 are not affected by the changes in the soil type specifications. Similar results (not shown) are obtained when the selected soil types are shifted by -1 in their numbers, that is, from type 7 to type 6 (clay loam) for the upper layer and from type 9 to type 8 (sandy clay) in the lower layer. Thus, the retrievals are moderately sensitive to errors in the soil type specifications.

5. Conclusions

In this paper, the relationship between soil temperature and soil water content is shown to be useful for retrieving soil water content from soil temperature measurements. In particular, a simple statistical method is developed to retrieve the bulk soil water content averaged through the depths of the three temperature measurements. The method is applied to 50 Oklahoma Mesonet stations where soil temperature was measured at three depths during the 10 dry summer days from 11 to 20 June in 1995. The retrieved bulk soil water contents are verified against the direct measurements for one pair of nearly collocated Mesonet #38 and ARM #12 stations and also compared with the retrievals of the previous method of ZX99 at 14 enhanced Oklahoma Mesonet stations. The results show that the current method gives more persistent retrievals than the previous method, and the retrieved persistence agrees with the fact that there was no precipitation over the entire Oklahoma Mesonet area during the 10 days.

The method is also applied to Oklahoma Mesonet soil temperature data (at 5, 25, 60, and 75 cm) from the Norman site during the dry period of 20–30 July 1998 and the wet period of 1–31 July 2000. The retrieved soil water contents are compared with the direct measurements, and the rms differences are smaller than the soil water observation error ($0.05\ \mathrm{m}^3\ \mathrm{m}^{-3}$). The rms differences are generally larger for the wet period than for the dry period, so the retrievals have relatively large errors for the wet period. However, the relative errors (with respect to the observed soil water contents) are about the same for the two periods. Based on the correlation coefficient (see R values in column 5 of Table 1), the linear regression fitting is significant (within confidence level $\geqslant$95%) only in the upper layer (from –5 to –60 cm). In the lower layer (from –25 to –75 cm), the diurnal variations of soil temperature are found to be too weak to have the required statistical significance. Although not shown in this paper, the method is also tested with data collected from the Norman site for all of July 2001 and July 2002. The test results are similar to those reported in section 4c for July 2000.

The bias part of the equation error is assessed by the regression intercept γ, while γ is found to be small in the upper layer but not so in the lower layer (see column 2 of Table 1). The random part of the equation error is assessed by the residual σ of the linear regression fitting, while σ is found to be large, especially in the lower layer (see column 4 of Table 1). Limited by the linear regression, the current method and its extension (see section 2.3) can only model the bias part of the equation error. While the bias part is loosely modeled by γ, the random part is large and not modeled. The empirical formulation of soil thermal conductivity (McCumber and Pielke, 1981) used in this paper neglects the effect of heat transport due to vapor and/or liquid water movement in the soil (Cahill and Parlange, 1988). According to Peters-Lidard et al. (1998), this empirical formulation tends to overestimate (underestimate) the soil thermal conductivity during a wet (dry) period. The could cause an error in the equation. This error is not assessed but could be a small part of the equation error assessed above.

The method is limited by its dependence on the diurnal variations of soil temperature. Obviously, the method cannot be applied to a soil layer in which there is no detectable daily temperature variation. Nevertheless, the results in this paper show that the method can work even when the daily temperature variations are significant ($\geqslant 2$ K) only at one depth (mostly in the upper layer). As explained in section 2.4, the retrieved soil water contents are not affected by observation bias errors (mainly calibrations errors within ± 0.5 K), but they are moderately sensitive to random errors (between ± 0.1 K) in the soil temperature observations (see column 1 of Table 2 versus Table 1) and moderately sensitive to errors in the soil type specifications. With the current method, the random part of the observation error is mixed into the random part of the equation error in the linear regression equation, so the random error is not only large but also cannot be modeled. Modeling the random part of the equation error statistically (essentially the error covariance) requires an advanced method (such as the adaptive Kalman filter). To this end, the study presented in this paper is intended to serve as an introduction to the problem of retrieving soil water contents from soil temperature observations.

Acknowledgments. The authors are thankful to Todd Crawford and Jerry Brotzge for their help with the data from the Norman site, and to Dr. Jeff Basara and anonymous reviewers for their comments and suggestions that improved the presentation of the results. This work was supported by the ONR Grant N000140310822 and the DOE-ARM Program through Battelle PNL Contract 144880-A-Q1 to CIMMS, University of Oklahoma.

REFERENCES

Al Nakshabandi, G., and H. Kohnke, 1965: Thermal conductivity and diffusivity of soils as related to moisture tension and other physical properties. *Agric. Meteor.*, **2**, 271–279.

Basara, J. B., and T. M. Crawford, 2000: Improved installation procedures for deep-layer soil moisture measurements. *J. Atmos. Oceanic Technol.*, **17**, 879–884.

Cahill, A. T., and M. B. Parlange, 1988: On water vapor transport in field soils. *Water Resour. Res.*, **34**, 731–739.

Campbell, G., 1985: *Soil Physics With Basic-transport Models for Soil-plant Systems*. Elsevier, New York, 150pp.

Clapp, R. B., and G. M. Hornberger, 1978: Empirical equations for some soil hydraulic properties. *Water Resour. Res.,* **14**, 601–604.

Hillel, D., 1980: *Fundamentals of Soil Physics.* Academic Press, 413pp.

Koorevaar, P., G. Menelik, and C. Dirksen, 1983: *Elements of Soil Physics*. Elsevier, New York, 228pp.

McCumber, M. C., and R. A. Pielke, 1981: Simulation of effects of surface fluxes of heat and moisture in a mesoscale numerical model. *J. Geophys. Res.*, **86**(C10), 9929–9938.

Peters-Lidard, C. D., E. Blackburn, X. Liang, and E. F. Wood, 1998: The effect of soil thermal conductivity parameterization on surface energy fluxes and temperatures. *J. Atmos. Sci.*, **55**, 1209–1224.

Pielke, R. A., 1984: *Mesoscale Meteorological Modeling.* Academic Press, Orlando, Florida, 612pp.

Reece, C. F., 1996: Evaluation of a line heat dissipation sensor for measuring soil matric potential. *Soil Sci. Soc. Amer. J.*, **60**, 1022-1028.

Schneider, J. M., D. K. Fisher, R. L. Elliott, G. O. Browen, and C. P. Bahrmann, 2003: Spatiotemporal variation in soil water: First results from the ARM SGP CART network. *Journal of Hydrometeorology*, **4**, 106–120.

Xu, Q., B. Zhou, S. D. Burk, and E. H. Barker, 1999: An air-soil layer coupled scheme for computing surface heat fluxes. *J. Appl. Meteor.*, **38**, 211–223.

Zhou, B., and Q. Xu, 1999: Computing surface fluxes from mesonet data. *J. Appl. Meteor.*, **38**, 1370–1383.