Correcting Temperature Sensitivity of ECH$_2$O Soil Moisture Sensors Application Note

Doug Cobos and Colin Campbell
Decagon Devices, Inc.
Pullman, WA

Introduction

Onset adds smart sensor adapters to the ECH$_2$O probes so that they are plug-and-play compatible with HOBO® U30 stations and HOBO Micro Stations. The S-SMC-M005 incorporates the EC-5 and the S-SMD-M005 incorporates the 10HS. Onset also offers the W-SMC HOBOnode wireless soil moisture sensor, which uses the ECH$_2$O EC-5 and provides data in the same format as the S-SMC-M005.

In many natural and engineered soils, the output of ECH$_2$O soil moisture sensors is sensitive to variations in the soil temperature. The temperature sensitivity is not caused by the ECH$_2$O sensors themselves which are almost perfectly insensitive to temperature changes, but rather the electrical characteristics of the soil, which can be quite sensitive to temperature changes. The ECH$_2$O sensors measure volumetric water content (VWC) by measuring the dielectric permittivity (ε) of the bulk soil. ε in the soil is a complex quantity with both real (ε') and imaginary (ε'') components. ε' is the real dielectric permittivity of the soil constituents, and has a negative correlation with temperature. ε'' is related to dielectric losses, and more importantly electrical conduction through the soil. The ability of a soil to conduct electrical current or electrical conductivity (EC) is related to VWC and to the amount of free ions in the soil (solute content). The EC of a soil has a strong positive correlation to temperature. The opposing temperature sensitivities of the real and imaginary components of the dielectric permittivity can be thought of as opposing forces in the soil. In some soils, ε' dominates, and an increase in temperature causes a decrease in the VWC measured by the ECH$_2$O sensor. In other soils, ε'' dominates, and an increase in temperature causes an increase in the VWC measured by the ECH$_2$O sensor. In some soils, the two components closely balance each other, and there is no apparent temperature sensitivity in the VWC measurement. Because of these complex interactions, it is impossible to determine a generic correction factor for temperature that can be applied to all soils.

The older ECH$_2$O sensors (EC-10 and EC-20) operate at a low measurement frequency, and are more strongly affected by ε'', meaning that there is often a positive relationship between temperature and measured VWC. The new generation of ECH$_2$O sensors (EC-5 and 10HS) operate at a much higher measurement frequency, and are much less affected by salts in the soil (ε''). With these sensors, the temperature sensitivity is generally small, and can be either negative or positive. It should be noted that any ECH$_2$O sensor buried at a depth of more than about 15 cm in the soil, or under a full vegetative canopy, will have little or no noticeable temperature sensitivity due to the very small diurnal fluctuations in soil temperature at these depths or under a closed canopy. The strategies described below for correcting the temperature sensitivity of ECH$_2$O sensors are meant primarily for users who have sensors placed in the top 15 cm of the soil profile under a bare surface, or whose sensors otherwise undergo strong temperature cycling.Strategy 1: Multiple regression analysis

If temperature data are available at the same location as the ECH$_2$O sensor, then a multiple regression strategy can be used to relate the true VWC to the measured VWC and temperature data. The general goal is to construct a mathematical model (equation) of the form:

$$VWC_{corrected} = C1 * VWC_{meas} + C2 * T_{soil} + C3 \quad (1)$$

Where VWC_{meas} is the VWC measured by the ECH$_2$O sensor, T_{soil} is the soil temperature at the location of the sensor, and C1 – C3 are empirical coefficients determined by multiple regression on field data. The steps below describe how to use the regression tool in Excel and co-located ECH$_2$O sensor and temperature data to determine the values of C1 – C3 for a particular soil.

Requirements:

The requirement for this method is co-located temperature and ECH$_2$O sensor output data in the soil that you wish to construct the correction model for. The data must have at least two (preferably three or more) 24 hour periods with no rain or irrigation events. For maximum effectiveness, the chosen 24 hour periods should have different average VWC that span the range from wet soil to dry soil (i.e. one 24 hour period with dry soil, one with moderately wet soil, and one with quite wet soil).

Method:

1. Identify three or more 24 hour periods to include in the data analysis (see Figure 1 for example). These periods should meet the following qualifiers:
    1.1. No precipitation or irrigation in each 24 hour period
    1.2. Group of 24 hour periods spans the range of expected VWC
    1.3. Temperatures at beginning and end of 24 hour period are comparable
    1.4. No anomalies present in the selected data

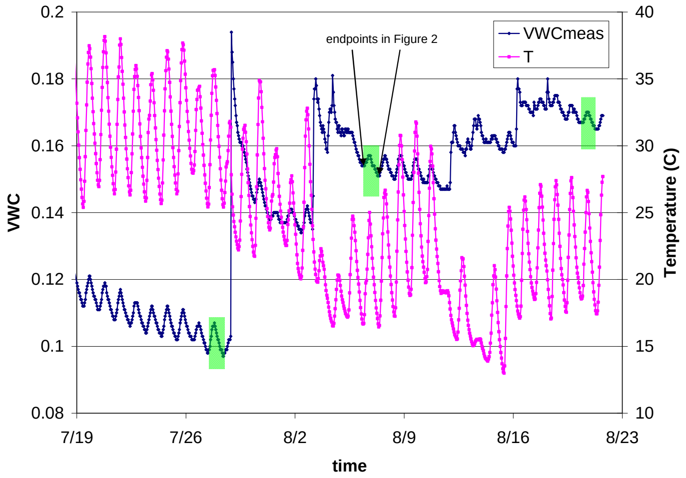


Figure 1. Example of thee 24 hour periods chosen from a data set (shown under shaded boxes). The endpoints used for the linear interpolation example data shown in Figure 2 are indicated.

2. Interpolate VWC data between endpoints. To determine the non-temperature affected VWC of the soil during each 24 hour period, constant drainage/evaporation is assumed over the period, and VWC is interpolated as a straight line between the two endpoints of each 24 hour period. This is easily accomplished in Excel by plotting the VWC at the two endpoints on the Y-axis against time on the X-axis. Then, use the trendline function to draw the regression line between them (see Figure 2). The resulting linear equation is then applied to all of the time data in that 24 hour period to interpolate between the two endpoints. Often it is easiest to assign an arbitrary time scale to the data for this part of the correction instead of the time stamp your data acquisition system outputs. For example, if you are collecting data at 30 minute intervals, you would assign X-axis values of 1 to 49 during the interpolation routine, or values of 1 to 25 for hourly data (column A in Table 2). Note that this process must be repeated for each of the 24 hour periods.

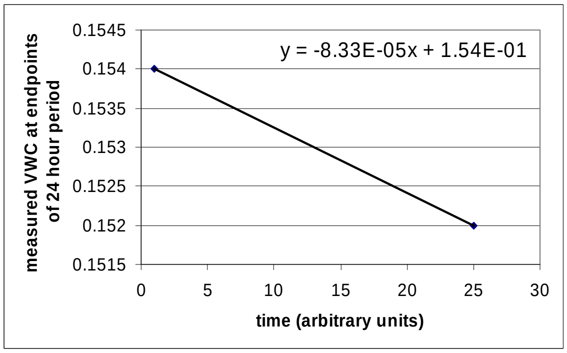


Figure 2. Linear interpolation between two endpoints of one 24 hour period (indicated in Figure 1). The X-axis is column A from Table 2, and the Y-axis is column E from Table 2. The linear equation was generated using the trendline function.

Table 2. Example of linear interpolation for a single 24 hour period. Note that the formula shown in cell F3 comes from the linear regression equation from Figure 2, and is applied to cells F3 to F25. Also note the arbitrary time scale in column A. This is often much easier to work with than the timestamp data from many data acquisition systems.

	A	B	C	D	E	F
1	time scale for interpolation	time	VWCmeas	T	VWC for interpolation	interpolated VWC
2	1	8/6/2006 11:00	0.154	17.36	0.154	0.154
3	2	8/6/2006 12:00	0.155	18.93		=-0.0000833*A3+0.154
4	3	8/6/2006 13:00	0.155	19.74		0.154
5	4	8/6/2006 14:00	0.155	19.93		0.154
6	5	8/6/2006 15:00	0.156	20.77		0.154
7	6	8/6/2006 16:00	0.156	21.87		0.154
8	7	8/6/2006 17:00	0.157	23.11		0.153
9	8	8/6/2006 18:00	0.157	24.41		0.153
10	9	8/6/2006 19:00	0.157	25.01		0.153
11	10	8/6/2006 20:00	0.157	24.14		0.153
12	11	8/6/2006 21:00	0.156	23.26		0.153
13	12	8/6/2006 22:00	0.155	22.44		0.153
14	13	8/6/2006 23:00	0.154	21.58		0.153
15	14	8/7/2006 0:00	0.154	20.83		0.153
16	15	8/7/2006 1:00	0.154	20.16		0.153
17	16	8/7/2006 2:00	0.154	19.61		0.153
18	17	8/7/2006 3:00	0.153	19.03		0.153
19	18	8/7/2006 4:00	0.153	18.42		0.153
20	19	8/7/2006 5:00	0.152	17.87		0.152
21	20	8/7/2006 6:00	0.152	17.47		0.152
22	21	8/7/2006 7:00	0.151	17.07		0.152
23	22	8/7/2006 8:00	0.151	16.68		0.152
24	23	8/7/2006 9:00	0.151	16.45		0.152
25	24	8/7/2006 10:00	0.151	16.57		0.152
26	25	8/7/2006 11:00	0.152	17.27	0.152	0.152

3. Combine all 24 hour periods and run a multiple regression. In Excel, this is easiest accomplished by pasting data from all 24 hour periods into the same columns (see Table 3). Then run the regression tool: Tools>Data Analysis>Regression. Select the interpolated VWC column from step 2 as the Y-variable (C2:C76 in Table 3), and measured VWC and Temperature as your X-axis variables (A2:B76 in Table 3). The regression tool will output several descriptive statistics and the multiple regression coefficients (see Table 4). The coefficients X Variable 1 and X Variable 2 are C1 and C2 in equation 1 above, and the Intercept is C3. Note that X Variable 1 is the coefficient for the first (leftmost) column of the X-axis variable range (in this case measured VWC).

Table 3. Combined data from three 24 hour periods, ready for multiple regression. With all 24 hour periods in the same columns, the multiple regression analysis will include data from all three days.

	A	B	C
1	VWCmeas	T	interpolated VWC
2	0.098	25.42	0.0979583
3	0.099	25.63	0.0979166
4	0.099	26.87	0.0978749

	A	B	C	
5	0.1	28.58	0.0978332	
6	0.102	30.43	0.0977915	
7	0.103	32.27	0.0977498	
8	0.105	33.92	0.0977081	
9	0.106	35.23	0.0976664	
10	0.106	35.64	0.0976247	
11	0.106	35.64	0.097583	
12	0.107	35.68	0.0975413	
13	0.106	35.25	0.0974996	
14	0.105	34.39	0.0974579	1st 24 hour period
15	0.104	33.39	0.0974162	
16	0.103	32.4	0.0973745	
17	0.102	31.51	0.0973328	
18	0.102	30.67	0.0972911	
19	0.101	29.95	0.0972494	
20	0.1	29.14	0.0972077	
21	0.099	28.28	0.097166	
22	0.099	27.43	0.0971243	
23	0.099	26.67	0.0970826	
24	0.098	26.03	0.0970409	
25	0.097	25.59	0.0969992	
26	0.097	25.41	0.0969575	
27	0.154	17.36	0.1539167	
28	0.155	18.93	0.1538334	
29	0.155	19.74	0.1537501	
30	0.155	19.93	0.1536668	
31	0.156	20.77	0.1535835	
32	0.156	21.87	0.1535002	
33	0.157	23.11	0.1534169	
34	0.157	24.41	0.1533336	
35	0.157	25.01	0.1532503	
36	0.157	24.14	0.153167	
37	0.156	23.26	0.1530837	
38	0.155	22.44	0.1530004	2nd 24 hour period
39	0.154	21.58	0.1529171	
40	0.154	20.83	0.1528338	
41	0.154	20.16	0.1527505	
42	0.154	19.61	0.1526672	
43	0.153	19.03	0.1525839	
44	0.153	18.42	0.1525006	
45	0.152	17.87	0.1524173	
46	0.152	17.47	0.152334	
47	0.151	17.07	0.1522507	
48	0.151	16.68	0.1521674	
49	0.151	16.45	0.1520841	
50	0.151	16.57	0.1520008	
51	0.152	17.27	0.1519175	
52	0.167	19.54	0.1669583	
53	0.168	21.14	0.1669166	
54	0.168	22.31	0.1668749	
55	0.169	23.54	0.1668332	
56	0.169	25	0.1667915	
57	0.17	26.17	0.1667498	
58	0.17	26.65	0.1667081	
59	0.17	26.51	0.1666664	
60	0.17	25.83	0.1666247	
61	0.169	24.91	0.166583	
62	0.169	23.84	0.1665413	3rd 24 hour period
63	0.168	22.84	0.1664996	
64	0.168	21.99	0.1664579	
65	0.167	21.33	0.1664162	
66	0.167	20.54	0.1663745	
67	0.166	19.76	0.1663328	
68	0.166	19.07	0.1662911	
69	0.166	18.51	0.1662494	
70	0.165	18.09	0.1662077	
71	0.165	17.66	0.166166	
72	0.165	17.44	0.1661243	
73	0.165	17.42	0.1660826	
74	0.165	17.65	0.1660409	
75	0.165	18.3	0.1659992	
76	0.166	19.56	0.1659575	

Table 4. Example output from Excel regression function. Shaded cells contain the information important for this analysis.

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.99954768
R Square	0.999095565
Adjusted R Square	0.999070442
Standard Error	0.000916358
Observations	75

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	2	0.06678705	0.033393525	39767.8662	2.6889E-110
Residual	72	6.04592E-05	8.39711E-07		
Total	74	0.066847509			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	**0.016239203**	**0.001353677**	**11.99636264**	**7.70761E-19**	**0.013540697**	**0.018937709**	**0.013540697**	**0.018937709**
X Variable 1	**0.965494597**	**0.005510463**	**175.2111411**	**1.83299E-96**	**0.954509689**	**0.976479505**	**0.954509689**	**0.976479505**
X Variable 2	**-0.000562114**	**2.79704E-05**	**-20.09677639**	**3.02271E-31**	**-0.000617872**	**-0.000506356**	**-0.000617872**	**-0.000506356**

4. Construct a VWC correction model. Use the coefficients determined from the multiple regression in part 3 to construct a VWC correction model from equation 1. In the case of the example data, the model would be

   $VWC_{corrected} = 0.9655 * VWC_{meas} - 5.621 \times 10^{-4} * T_{soil} + 1.624 \times 10^{-2}$

5. Apply the model to your raw data set, yielding corrected VWC. The temperature dependency in the corrected VWC data should be greatly reduced. In some instances, there will still be some apparent temperature dependence. Often the results can be improved by selecting data to input into the correction algorithm that more fully span the range of VWC encountered in the soil.

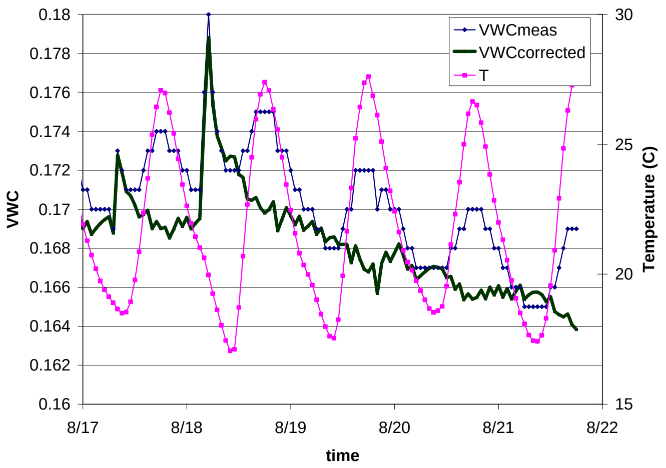


Figure 3. Example of temperature-corrected data. Notice that with this method the VWC dynamics are preserved, but the temperature dependency is greatly reduced compared to the uncorrected data.

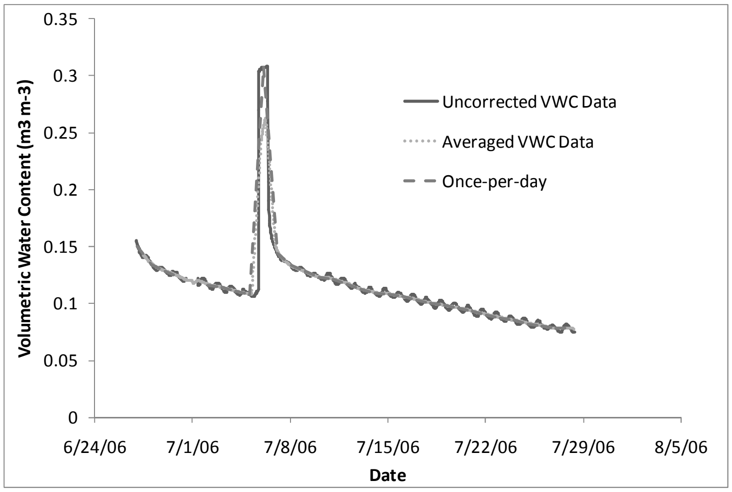
Strategy 2: Once-per-day measurements or averaging

Soil moisture is reasonably conservative and does not change much in a day unless a precipitation or irrigation event occurs. So, if no temperature information is available, or if a simplified smoothing technique is preferred, using once-per-day water content measurement or designing a simple averaging technique is useful. Of course, this would wash out any hourly changes that might be of interest, but in our experience, we generally see very little temperature-related fluctuations in the water content data when there are multiple rain or irrigation events during a day so these data would not need correction.

Implementing a once-per-day strategy involves selecting a time to use, culling and sorting the data, and checking to make sure the results fit the original data satisfactorily. We have chosen early morning hours (e.g. 2 AM) where temperature is near its minimum as our daily measurement. This reduces the likelihood of variation in the soil moisture data from changing daily high temperatures. To sort a dataset for 2 AM data, we use an "if" statement in Excel that only puts data into an adjoining column when it is 2 AM (Table 4). Once you have set up the formula for your worksheet, copy and paste so it covers all the date values and any water content columns you want to include. Next, use the sort command to collapse all the spaces in the data and sort your data by date. You should be left with a fraction of your original data that you can now graph against the full data to see the fit. Figure 4 shows that this method does a good job smoothing the data. Obviously this technique will not preserve the high frequency soil moisture dynamics, but the overall trends in soil moisture will remain.

Table 4. Example of Excel "If" statement to cull once-per-day date and water content values

	Q	R	S	T
5	6/27/06 12:00 AM	0.141	0.141	=IF(($Q5-TRUNC($Q5))/(120/(60*24))<1.1,IF(($Q5-TRUNC($Q5))/(120/(60*24))>0.9,Q5,""),"")



Data averaging is another way that you can smooth temperature cycling in the water content data. The advantage of this technique is that you can use the entire dataset instead of removing all but one reading per day. Implementing this technique is straightforward; simply make a new column for smoothed water content data and select the "Average()" function in Excel. For the data range, you must average an entire 24 hour period to remove the diurnal temperature cycle so the average function range in any given cell must include data from the past 12 hour and the next 12 hours. Data from the averaging are shown in Figure 4. From our analysis, this averaging technique did not appear to give better results than the once-per-day technique and, not surprisingly, tended to smooth large changes in water content.