Electrical Conductivity Characterizations and Development Using SM's Hindi-TRACE TDR System.

By Keith

Internal Report

1. Introduction

The waveform data measured with the HandiTrace TDR contains a lot of information about the soil, including soil moisture, salinity, and possibly texture and bulk density. The HandiTRACE currently has soil moisture calibrations in the form of look up tables rather than polynomials. For the metallic waveguides, the calibrations are based in part on the Topp equation for soil moistures up to about 50% and then draws a straight line to 100% to represent a dielectric constant of 80. For the coated tines, a lookup tables used are linear for the entire range.

The waveform is a plot with millivolts on the y axis and a time index on the x axis. The time index multiplied by the measurement interval time in ps; (Time index X Interval = Waveform Time in ps). The apparent permittivity K_A is determined by the travel time of the reflected pulse using equation [1]. A 5-volt pulse in the form of a single square wave is sent from the "Backpack" signal processing unit down the cable and out the tines where the signal is reflected back to the Backpack. The Backpack logs the voltage as a function of time for 1200 points with a selectable interval. The time of travel can be determined from the waveform for the determination of soil moisture using equations [1] and a lookup table based on the Topp Equation, Equation [2]. The electrical conductivity can be determined from the shape of the waveform at certain points.

$$t = \frac{2L\sqrt{K_A}}{c} \qquad [1]$$

The mathematical relationship between the time of travel, t, and the apparent permittivity, K_A, is shown in [1] where L is the length of the waveguide, c, is the speed of light.

The Topp equation is a widely used general purpose mineral soil moisture calibration used by most TDR manufactures, Equation [2].

$$\theta = -5.3x10^{-2} + 2.92x10^{-2}K_A - 5.5x10^{-4}K_A^2 - 4.3\ x\ 10^{-6}K_A^3 \qquad [2]$$

2. Bulk Electric Conductivity Calculation from the Waveform

Figure [1], shows the waveform and the features used to estimate bulk EC.
The t in equation [1] is t_{start} minus $t_{finish.}$ The slope term, S_f, is the maximum slope after V_3 using a center looking 20-point moving slope. Both t_{start} and t_{finish} were determined by the HandiTRACE algorithms in the Android operating system most of the time. For some samples however, the start and finish time had to be adjusted manually (a.k.a moving Pluto's balls) when the onboard algorithms missed the targets.

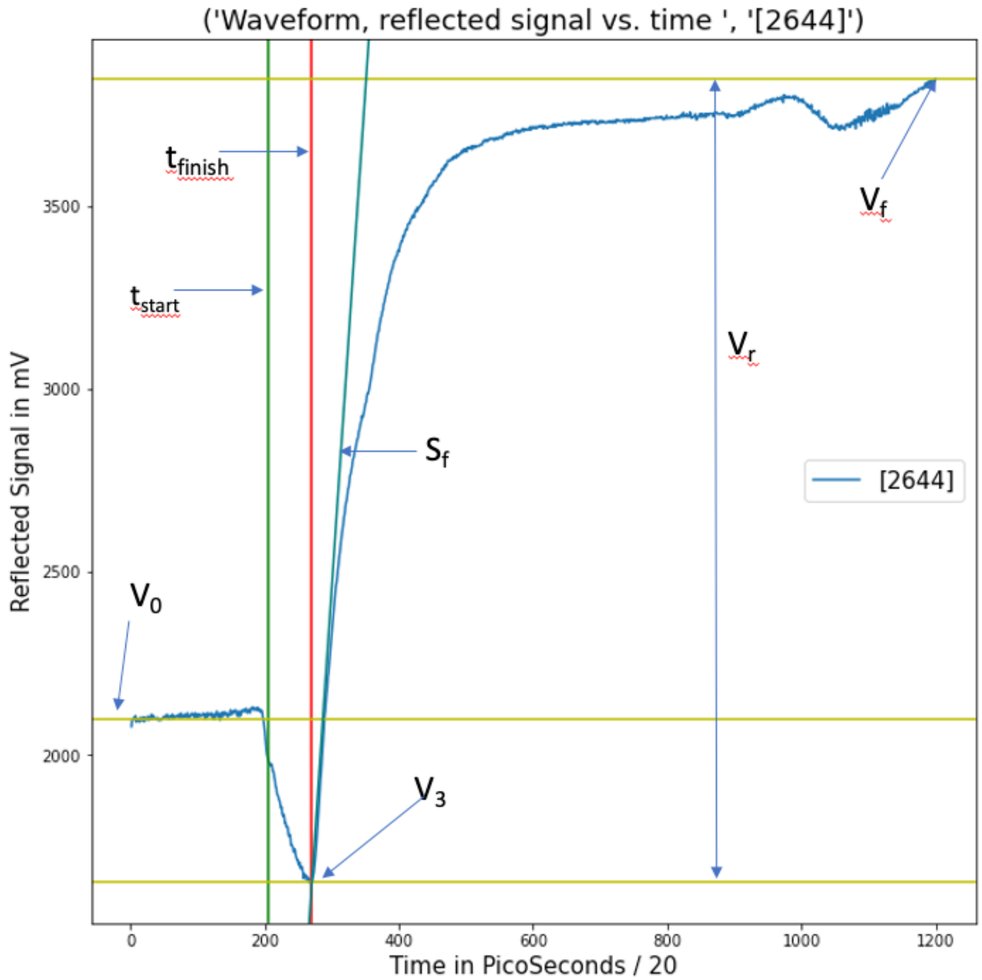

Figure [1], Waveform components

Variables	Definitions
t_{finish}	Time when the signal returns
t_{start}	Time when the signal enters the soil
V_0	Voltage at the beginning of the waveform
V_3	Voltage at t_{finish}
S_f	Maximum rolling slope after V_3
V_f	Maximum voltage after V_3
V_r	$V_f - V_3$
Tp	Output to the Topp EC model
L	Waveguide length in meters
σ_m	Topp output mapped EC
σ_{bulk}	Bulk EC

Table [1], Variable abbreviations and definitions.

Both the Topp and Dalton EC models was evaluated but the Topp model shown in equation [3] was used because it had a better fit:

$$T_p = \sqrt{K_A} \Big/ 120\pi L \ln[\frac{V_3(2V_0-V_3)}{V_0(V_r-V_3)}] \qquad [3]$$

The T_p term in equation [3] is then fit to a parabola to better map it the specific HandiTRACE measurement network.

$$\sigma_m = AT_p^2 + BT_p + C \qquad [4]$$

The mapped EC (σ_m) is then fit using a linear regression for better accuracy.

$$\sigma_{bulk} = b + m_{v0}V_0 + m_{v3}V_3 + m_{vr}V_r + m_{vf}V_f + m_{Ka}Ka + m_{sf}S_f + m_m\sigma_m \qquad [5]$$

The final bulk EC is then calculated in four steps:

1) Pull the variables $V_{0,}$ V_3, $V_{r,}$ V_f, Sf and K_a from the waveform.
2) Calculate Tp using equation [3]
3) Calculate the mapped Topp output using equation [4]
4) Calculate the bulk EC from Equation [5]

Steps to production are in section 6 and the coefficients to equations [4] and [5] are in tables [2] and [3]. Table [6.c] shows the accuracy which is about +/- 0.01 S/m from 0 to 0.6 S/m.

Waveguide	A	B	C
BUR 7.8	0.302	0.2549	0.0324
BUR 20	4.4553	-0.1841	0.0127
FLD 40	8.8467	-0.3415	0.0125
FCT 40	-9.5701	5.3563	0.0048

Table [2], Fitting parameters for equation [4]

Waveguide	MKa	Mvo	M3	Mvf	Mvr	Msf	Mm	b
BUR 7.8	-0.000187	0.000011	-0.00004	-0.000033	0.000007	-0.001074	0.928325	0.1851326
BUR 20	-0.007577	0.000498	-0.000502	-0.000623	-0.000121	0.04823	0.694863	1.390367
FLD 40	0.000698	-0.000025	0.000039	0.000112	0.000073	-0.013029	1.15477	-0.33951
FCT 40	-0.000657	0.000044	0.000014	0.000081	0.000067	-0.00019	1.17164	-0.5022252

Table [3], Fitting parameters for Equation [5]

3. Test Solutions

Eighteen Reference solutions were prepared in 40 L drums. The salt used was reagent grade potassium chloride and sodium chloride. Four solutions were made from medium grade sand which includes, dry soil, saturated with no salinity, saturation with medium salinity and saturation with high salinity. Because we only have 12 barrels, six of the solution are aliquots, either a dilution with water or an addition of salt to a previously tested solution. The Reference EC was measured with a Cole-Parmer 0 to 200 mS/cm model EC meter. The meter was calibrated by InnoCal labs and is traceable to NIST. For the saturated soil samples, the EC was measured in the surface water and the Hilhorst equation was used to determine the bulk EC which was in turn validated with the HydraGO.

At 0.6 S/m with a mix of NaCl, KCL and sand, this media is a corrosive and abrasive to the metal of the tines and the base of the Slammer. Plastic containers with a rubber gasket are recommended for storing the test media. After each test, the Tine assemblies should be cleaned off with tap water. A good practice is to start with the lowest salinity concentration first and go to the highest. This prevents cross contamination of test media.

Solution ID	Remark	EC Ref 4/11/23	EC Ref 5/24/23	EC Ref 5/30/23	HydraGO Bulk EC Raw
Sol 1	Low EC aq, 17.225 g KCL/40 L	0.0796	0.0802		0.0815
VAT	Low EC aq, 10g KCL/40 L	0.0511	0.0511		0.057
VAT2	Higher EC aq, 100g KCL/40 L	0.424	0.425		0.385
Sol 2	Na K mix aq, 17.2 NaCl and 17.2 g KCl in 40L	0.1552	0.1568		0.16
Sol 3	Upper limit, 155g in 40L Mix Na and K		0.592		0.4
Sol 4	AQ from Vat 2, Add 30 g NaCl.		0.316		0.316
Sol 5	Tap Water	0.00735	0.00722		0.036
Sol 6	KCl AQ, 100 g KCl/ 40L	0.414	0.414		0.38
Sol 7	NaCl AQ, 100 g NaCl/40 L	0.437	0.437		0.38
Sol 8	AQ From Vat 1, Dilute after filling Sol 9		0.0284		0.032
Sol 9	In Soil, From VAT1		0.0594		0.055
Sol 10	In Soil, Sat Sand no Sal	0.0205	0.0199		0.02
Sol 11	In Soil, Dry Sand	0	0	0	0
Sol 12	In Soil, Fro Vat 2		0.413		0.372
Sol 13	Dilute Sol 6			0.258	0.25
Sol 14	Dilute Sol 7			0.282	0.27
Sol 15	Dilute Sol 4			0.241	0.232
Sol 16	Add 20 grams to Sol 1			0.1567	0.157

Table [4], Reference EC with HydraGO reading, All readings in S/m.

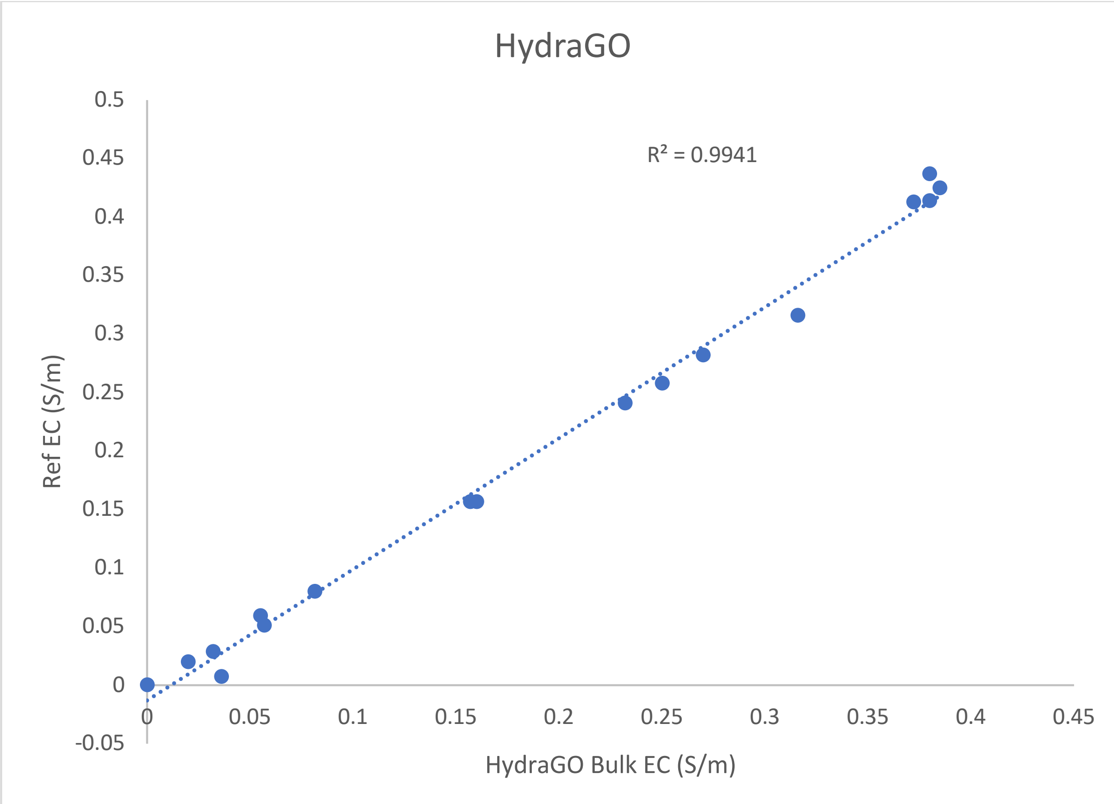

Figure [2], HydraGO readings with EC references. R^2 or 0.9941 and RMSE of +/- 0.02 S/m.

Note that the higher refence value of 0.592 S/m (Sol 3) was excluded from figure. If Sol 3 was included, the R^2 would be 0.96 with a RMSE of +/- 0.05 S/m.

4. Results, Accuracy and Model Development

Section 4 describes the development of Equations [3], [4], and [5], and the fitting parameters in Tables [2] and [3].

4.1 Evaluation of the Curve fitting of the Topp EC Model.

Equation [3] does not directly output a bulk EC value. It outputs a value that closely correlates to bulk EC (T_p), and in the HandiTRACES's case, this correlation is a parabola. A least squared technique was used to fit T_p to a bulk EC estimation, σ_m. The fitting parameters are in Table [3].

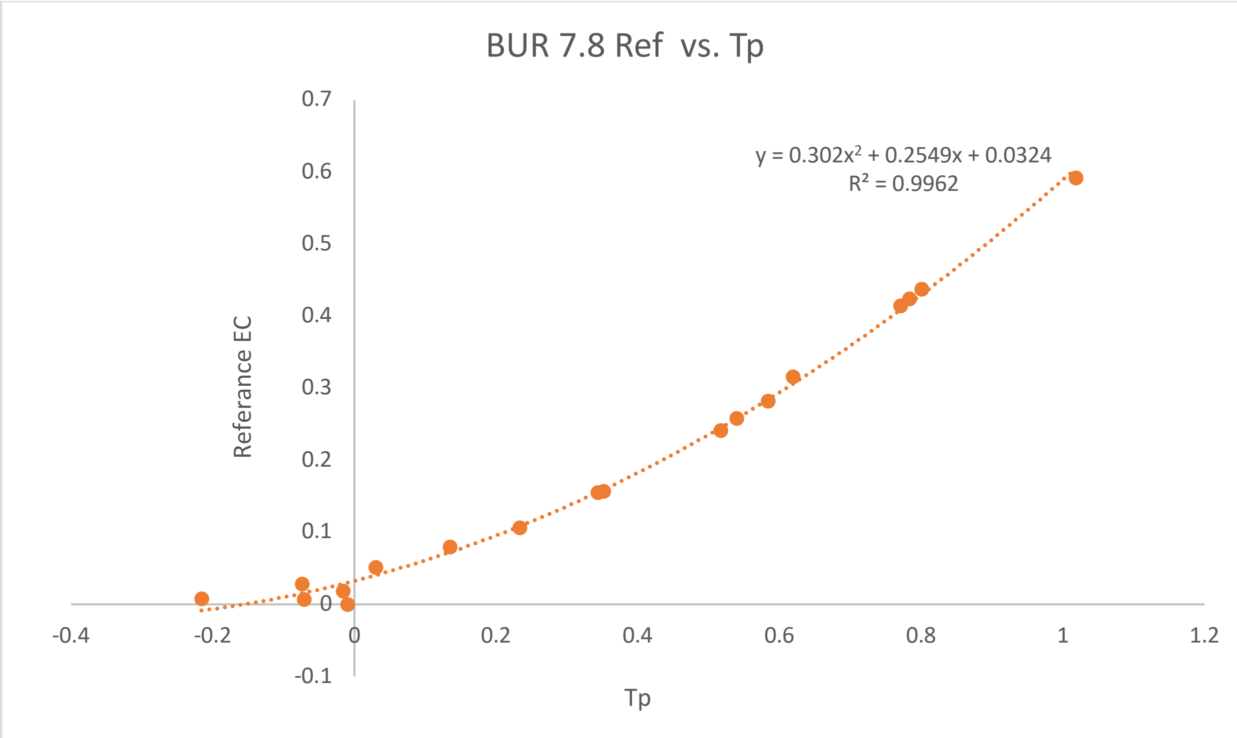

Figure [3], Output from equation [3], for BUR 7.8

Probe Type	R^2	RMSE
BUR 20	0.977	0.022
BUR8	0.9962	0.0108
FCT40	0.9909	0.0167
FLD40	0.9937	0.0121

Table [5], Accuracy of the parabolic curve fit to the Topp Model Output used in Equation [4]

4.2 Machine Learning Accuracy Boast

A linear regression model was used boast the accuracy from the results in Table [5]. In addition to parameters pulled from the waveform used in Equation [3], the ML model also included the output of Equation [4] plus the rise time (S_f) of the returned signal. The rise time is the rate of rise of the return signal which is a function of the impedance of the return signal and the bulk EC soil. It is modeled here as the maximum rolling slope after V_3 using center looking 20-unit points. Equation [5] is a linear regression that incorporates all of the input parameters. The coefficients and offsets for Equation [5], are listed in Table [3] and were determined using Scikit-learn machine learning python library.

4.3 Model Testing and Validation

With only 18 reference points, its important to test the model to make sure that it is not over fitting the reference data, particularly with the linear regression formula Equation [5]. This is called model training and validation. To do this, 5 of the 18 samples of the reference set were randomly removed and used as a validation set and the remaining 13 reference samples were used as a training set. A training model was constructed using only the 13 samples from the training set. The 5 samples from the validation set were modeled using the training model. This would test the training model on 5 samples that it had never seen before. The output of the training model on the 5 sample validation sets were then compared to the actual EC values. If the R^2 is over 0.90, than a new production model is made with all 18 references.

Probe Type	LR Train R^2	LR Train RMSE	Number of Samples
BUR 20	0.9985	0.005	13
BUR8	0.9998	0.00197	13
FCT40	0.9961	0.009	13
FLD40	0.9987	0.0045	13

Table [6.a], Model training performance

Probe Type	LR Validate R^2 with holdout	LR Validate Holdout RMSE	Test Holdout quantity
BUR 20	0.882	0.0974	5
BUR8	0.9948	0.0198	5
FCT40	0.9691	0.0485	5
FLD40	0.9385	0.0846	5

Table [6.b], Validation performance.

Probe Type	LR Final R^2	LR Final RMSE	Holdout quantity
BUR 20	0.985	0.0179	0
BUR8	0.99935	0.00454	0
FCT40	0.9934	0.01430	0
FLD40	0.9959	0.01303	0

Table [6.c], Final accuracy achieved.

Table [6] has the validation results with the holdouts. In general, the R^2 and the RMSE should increase from the validation step being the lowest, the training should be in the middle and the final should be the best. There are some exceptions caused by outliers. While there are no major outliers identified in the pair plots, we have minor outlier particularly with dry soil and the highest EC reference, Sol 3. Because the 5-holdout split is random, the minor outliers may randomly land in the training or the validation sets.

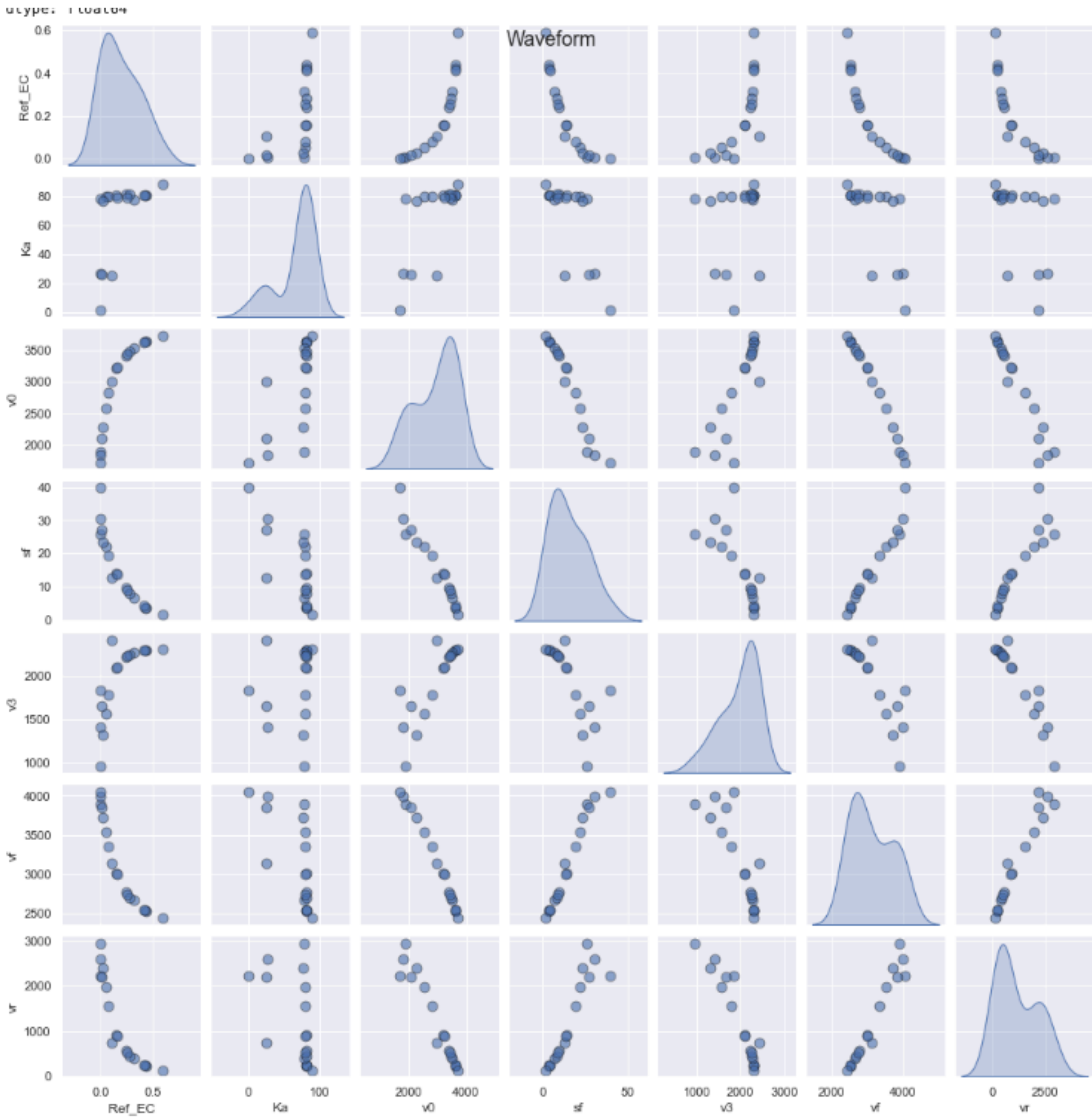

Figure [4], Pair plot for BUR 8 showing correlation between reference EC and Waveform parameters.

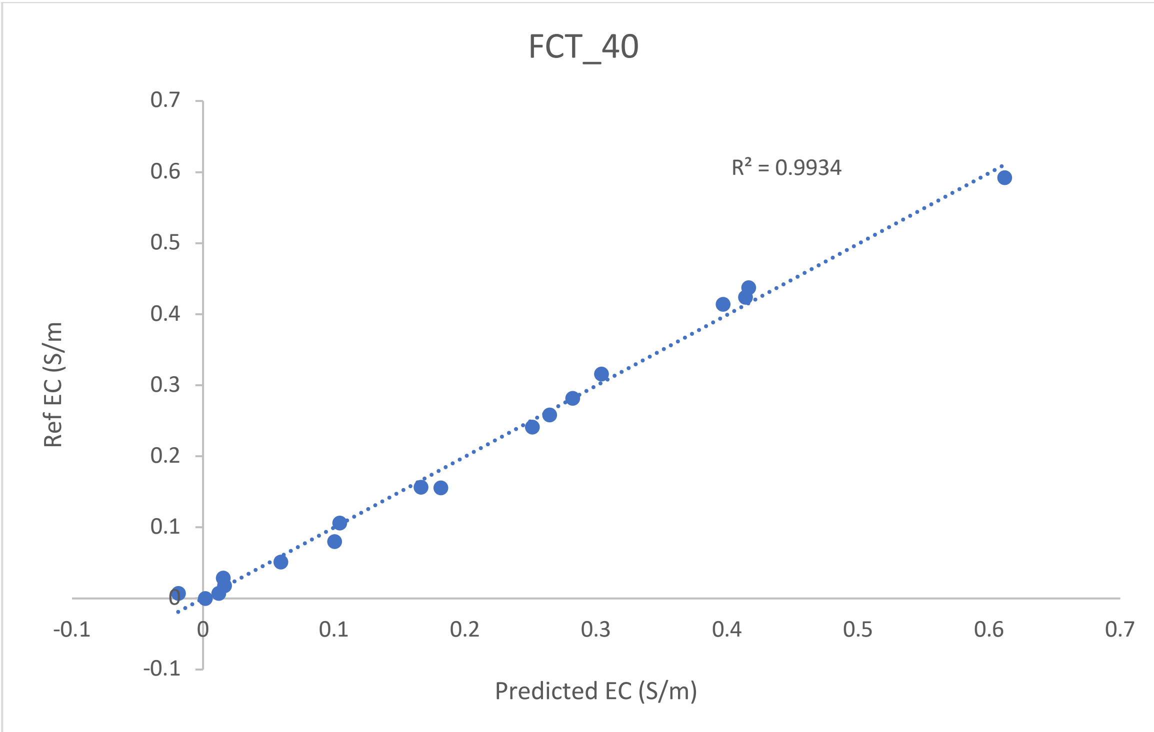

Figure [5.a], FCT 40 Final Model performance

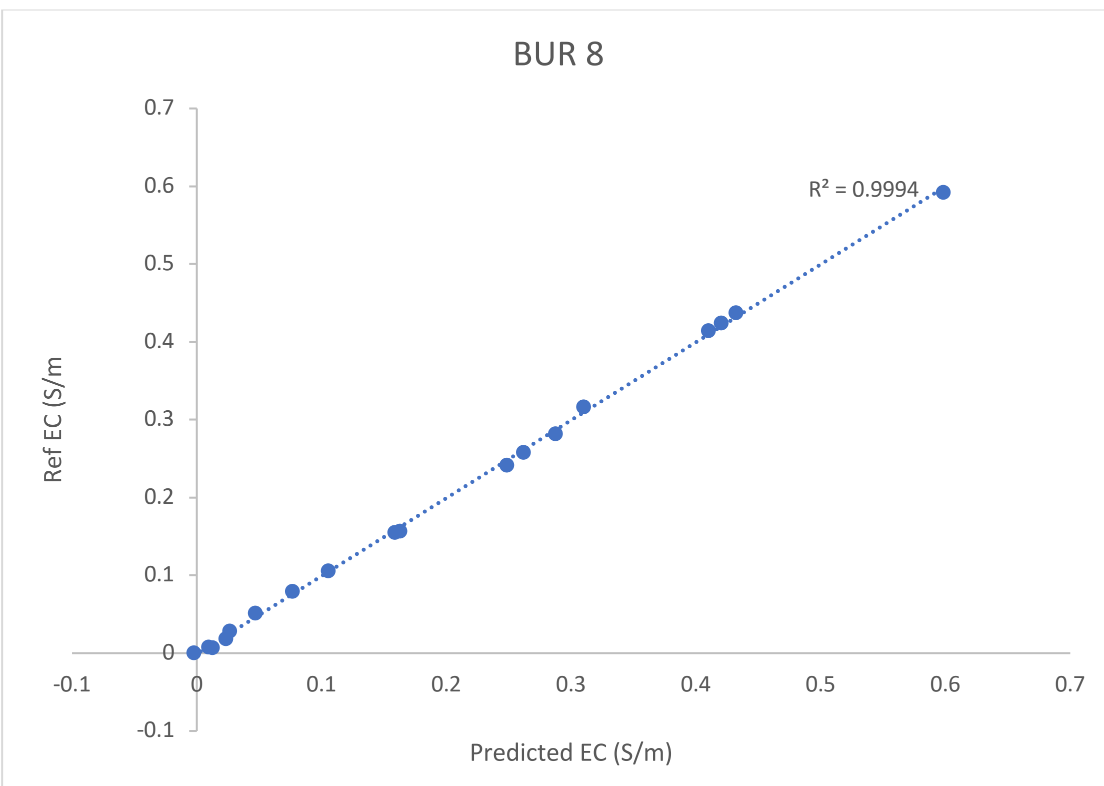

Figure [5.b], BUR 8 Final Model performance

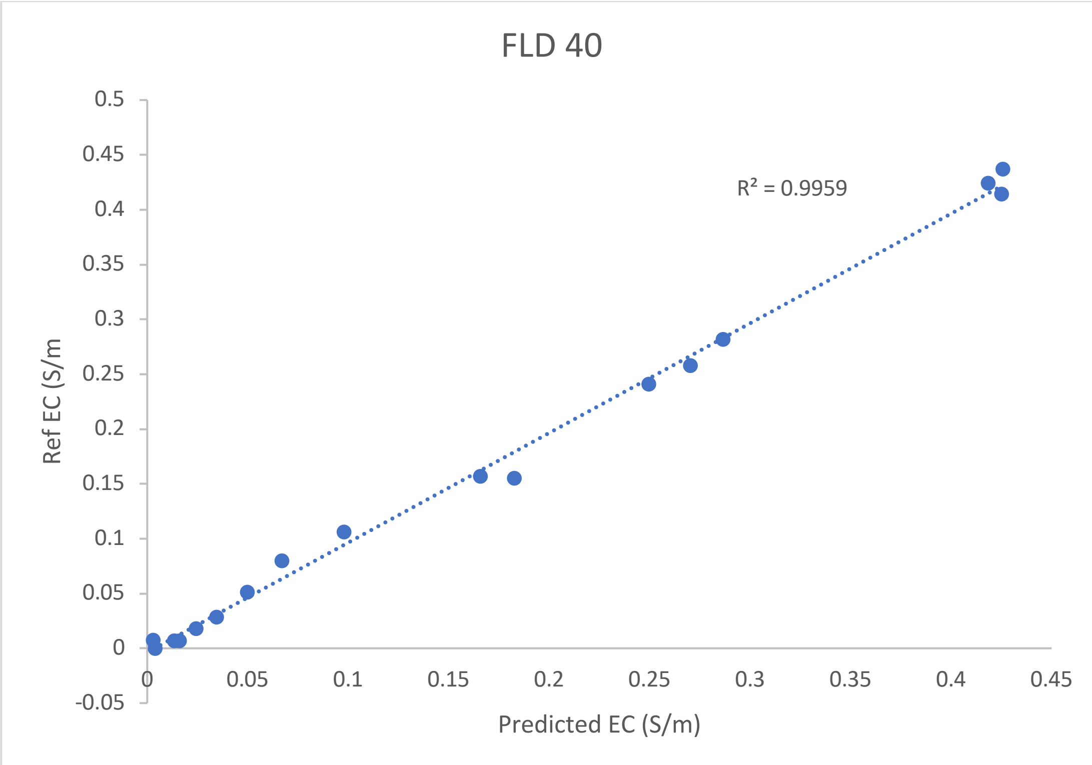

Figure [5.c], FLD 40 Final Model performance

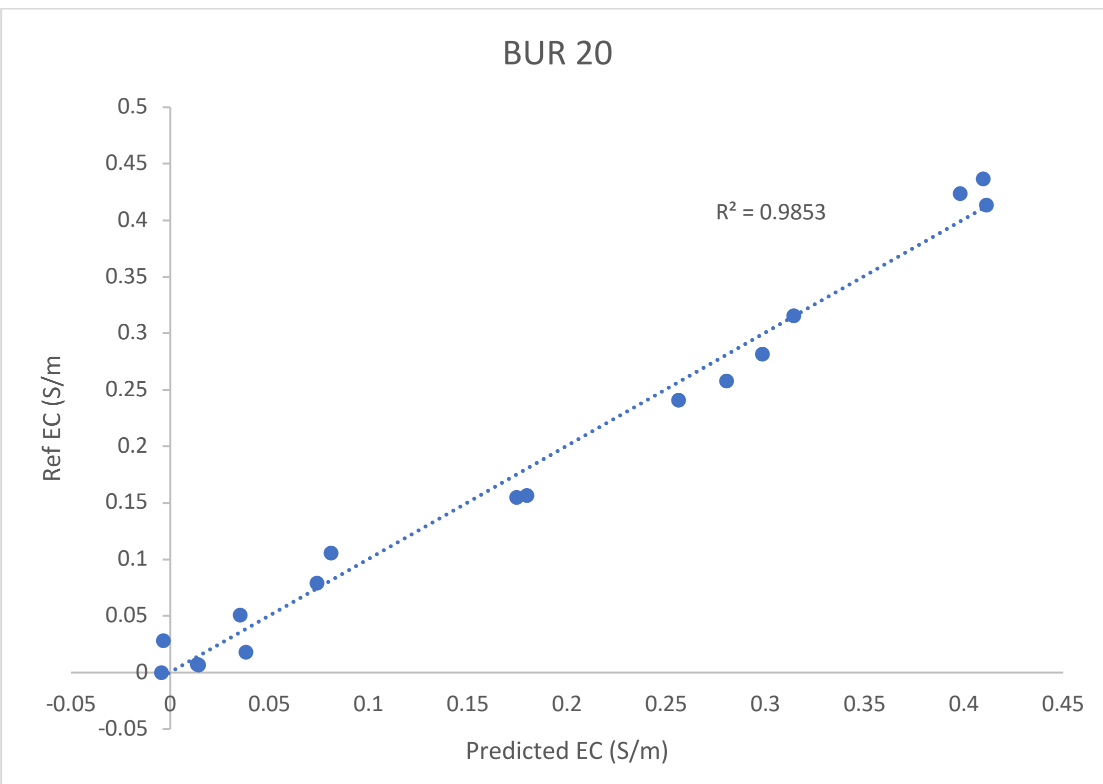

Figure [5.d] BUR 20 Final Model performance

4.6 Random Forest Model.

While a random forest can't be scripted in C and put on the HandiTRACE Android App without a ML Library, namely SciKit Learn, the same the train-test-split variables used for the linear regression were used in a random forest to see if further improvements can be made. This model and other ML models can be powerful tools for future developments with the HandiTrace technology such as tuning of the coated tines, bulk density, and textural classification from waveforms.

Probe Type	RF Train R^2	RF Train RMSE	Number of Samples
BUR 20	0.9929	0.0179	13
BUR8	0.9962	0.0116	13
FCT40	0.9973	0.01197	13
FLD40	0.9838	0.02	13

Table [7.a], RF Training

Probe Type	RF Validate R^2 with holdout	RF Validate Holdout RMSE	Test Holdout quantity
BUR 20	0.992	0.03	5
BUR8	0.9865	0.063	5
FCT40	0.9945	0.028	5
FLD40	0.961	0.044	5

Table [7.b], RF Validation

Probe Type	RF Final R^2	RF Final RMSE	Holdout quantity
BUR 20	0.997	0.0096	0
BUR8	0.9985	0.00654	0
FCT40	0.9976	0.00965	0
FLD40	0.9964	0.01134	0

Table [7.c], RF Final Model

The random forest was setup to have a maximum of 100 leaves per tree and the forest had 1000 trees.

5. Porewater EC

Porewater EC can be added as another EC parameter to The BUR 8, FLD 40 and the BUR 20. The porewater EC can't be used with any of the coated waveguides because the coating de-tunes the TDR to the point where the output K_A is no longer a true K_A value. The formula to use for porewater EC would be the Hilhorst equation.

$$\sigma_p = \frac{K_A \sigma_b}{\varepsilon_{rb} - \varepsilon_{rb\_O}} \qquad [6]$$

Where σ_p is the pore water EC, ε_{rp} is the real dielectric content of water (≈80), σ_b is the bulk EC measured with the HandiTRACE using Equation [5], and K_A is the apparent dielectric permittivity of the soil measure with the HandiTRACE using equation [1]. ε_{rb_O} is an offset, and 3.4 can be used as the offset for most inorganic soils and as the default.

6. Implementation of bulk and Porewater EC in the HandiTRACE

The formulas for calculating bulk EC are equations [3], [4], and [5] using values from tables [2] and [3]. To obtain the parameters from the waveform outlined in Table [1]; the HandiTRACE needs to be set at a scanning interval of 20 ps. While more precision can be achieved at a scan rate of 5 ps, the waveform is limited to 1200 points and with an interval of 5ps the V_f may get chopped off in wetter soils and in water.

With the exception of the Slammer (FLD 40) and coated Slammer (FCT 40), it is not possible to obtain the waveform parameters necessary for Equations [3] and [5] from any of the Field (FLD, FCT) or coated (BCT) waveguides. These waveforms are impaired from the coating which is double the thickness of the tines of the FLD and BUR waveguides. The waveform in the FLD 8 and FLD 20 are also distorted.

For the BUR 8, an electrical length of 7.8 cm is used for L. The BUR 8 tine is tapered and has a shorter electrical length. The length of 7.8 was determined from the waveform in distilled water at 20 degrees C matching to a dielectric constant of water at 78.8.

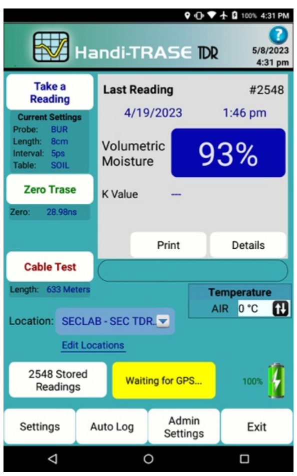

Figure [7], HandiTRACE App.

It is also important to have an accurate Start and Finish times. Not only are these times important for obtaining K_A (Used for both moisture and EC determination) but they are also important for determining T_f, S_f, V_3 and V_r. If the HandiTRACE Software doesn't automatically find T_f and T_s, manual adjustment using the App would be necessary to find the start and finish targets.

Bulk EC and porewater EC can be displayed in the App in a box under the soil moisture percent box. For Porewater EC, the offset can default to 3.4 but the user should be able to adjust it in the setting s section.

7. Bias and Error

The soil moisture measurement of the HandiTRACE is basically composed of three components; 1) the accurate measurement of voltage and time and the generation of the square pulse, 2) the process for determining a start and finish time, and 3) the soil moisture calibrations. Below is an assessment of the factors that could introduce error into the EC estimation as well as for soil moisture.

7.1 Reproducibility

All the data in this report was collected with a single set of waveguides and a single tablet with a single Backpack. The standard deviations between waveguides and waveforms are not known. The coated tines, 8 and 20 BCT are more than twice as fat as the noncoated version. The coating deforms the waveform but also may exhibit large variations.

Waveguide	Center Tine Diameter in mm
BUR 8	3.4
BCT 8	6.82
FCT 40	9.70
FLD 40	9.53
BCT 20	7.0
BUR 20	3.3

Table [8], Waveguide diameters

7.2 Sphere of influence

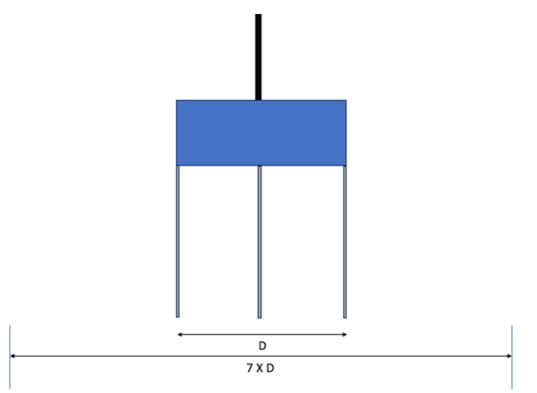

Figure [8], Sphere of influence of waveguide.

The sphere of influence of the waveguide is 7 times its diameter in the plain of the waveguide. It is approximately 5 times the diameter perpendicular to the waveguide. With this large measurement region, objects, and void spaces will affect the waveforms. Containers for testing must be wide enough to fit the 7X D diameter. Some qualitative tests were performed on the FLD 40 by taking readings in water with the tines next to the wall of the contain. These readings were taken both in and out of the waveguide plane and moved toward the middle of the vessel. Noticeable difference in waveforms were observed less than 3 X D regardless of tine orientation. Not much influenced was noticed beyond 3X D.

7.3 **Start and Finish Times**

The algorithms that determine the start and finish times need to be adjusted about 20% of the time for 8 and 20 BUR and FLD 40.

7.4 **Accuracy of hardware**

The voltage measurement, the interval time and the pulse generation components in the Backpack are made of quality components. A notch is place at the head of the waveguides to create a dip in the waveform. This allows the algorithm to find the start time.

Other than two Slammers, none of the FLD or Coated waveguides had interpretable waveforms. Figure [8] below shows large scatter in the pair plot of BCT 8 modeling parameters. Note the difference between figure [8] and figure [4] for the noncoated.

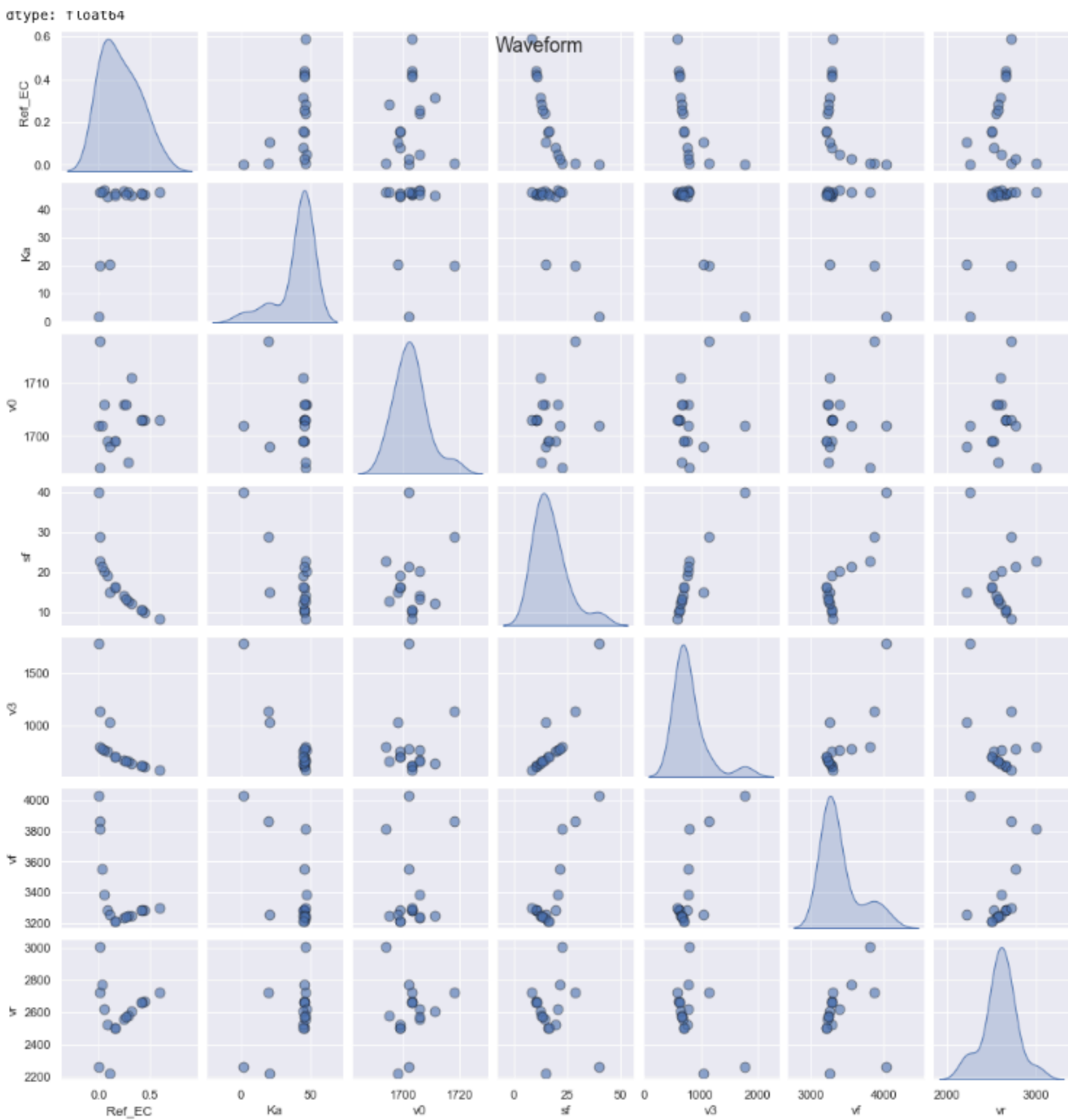

Figure [9], BCT 8pair plots.

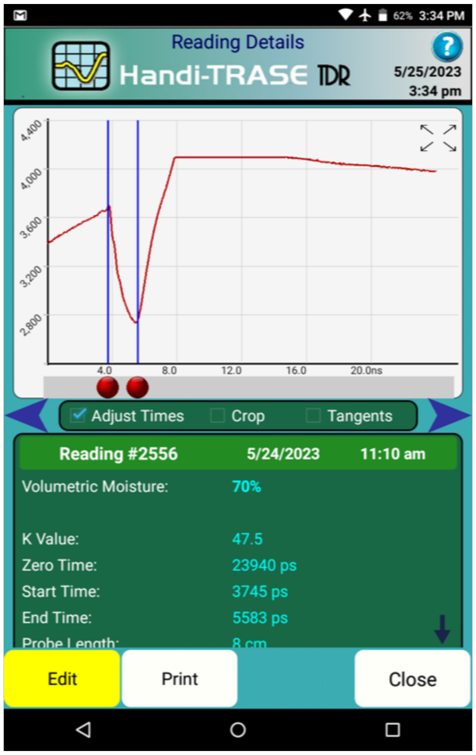

Figure [10], FLD 8

Figure [10] shows a FLD 8in saturated sand with no salinity. Saturation of this sand is between 39% and 40% as measured with a Tempe Cell, HydraProbe and the metallic waveguides. The permittivity is near 28 dielectric units. The FLD 8 waveguide generated a soil moisture value of 70% with a Ka of 47.5. The start and finish time appear to have been resolved by the software.

7.5 **Temperature**

All data was collected at ambient temperature. EC may be influenced by temperature. Temperature is not measured in the waveguides.

8. Conclusion

The waveform measured with the HandiTRACE TDR captures a lot of information about the soil. With 10 different waveguides, 134 waveforms were collected in 18 different solutions referenced with a NIST Traceable EC Meter. A series of mathematical formulas were developed to put EC on the HandiTRACE using the BUR 8, BUR 20, FLD 40, and FCT 40 waveguides.

The waveforms for the FCT 8, FCT20, BCT 8, BCT 20, FLD 8, and FLD 20 Waveguides are unsuitable for EC determination. The coating and the handheld unit severely deform the waveforms removing correlations with EC, (as shown in figure [9] and making the permittivity measurement un-referenceable to standards with known dielectric constants. The calibration is a table that fits a straight line with the distorted K_A.

Data files

Data files can be found at https://github.com/KeithBellingham/EC_HandiTRACE_TDR

Data File	Remark
EC Characterization and Development of HandiTRACE.pdf	This Report
HydraGO052323.csv	HydraGO Data
Rawwf061223.csv	Waveform export from HandiTrace App (Non-UTF8 encoding)
Trace_EC.xlsx	Spreadsheet workbooks
Waveform_raw053123.csv	Waveforms UTF-8, input for processing
WF_Processing.ipynb	Batch processing code
TDR_ANA_Trace.ipynb	Waveform analytical tool
Model_input.csv	Processing output, model input data
Model_BUR_7_8.ipynb	Modeling code for BUR 8
Model_FCT_40_7_8.ipynb	Modeling code for FCT 40
Model_FLD_40_7_8.ipynb	Modeling code for FLD 40
Model_BUR_20.ipynb	Modeling code for BUR 20

Python Libraries

Python was used to analyze data and curve fit the models. The libraries used were:

```
import pandas as pd
%matplotlib inline
import matplotlib.pyplot as plt
import seaborn as sns; sns.set()
import numpy as np
from sklearn.linear_model import LinearRegression
from sklearn.metrics import mean_absolute_error
from sklearn.model_selection import train_test_split
from sklearn.tree import DecisionTreeRegressor
from sklearn.ensemble import RandomForestRegressor
from sklearn.ensemble import RandomForestClassifier
from sklearn.linear_model import LinearRegression
from sklearn.metrics import r2_score
from sklearn.model_selection import cross_val_score
```

References

Dalton, F.N., Van Genuchten, M. Th. 1986. "The Time-Domain Reflectometery Method for Measuring Soil Water Content and Salinity." *Geoderma* 38: 237-250.

Ebrahimi-Birang, N. ,C. P. Maulé, W. A. Morley. 2006. "CALIBRATION OF A TDR INSTRUMENT FOR SIMULTANEOUS MEASUREMENTS OF SOIL WATER AND SOIL ELECTRICAL CONDUCTIVITY." *American Society of Agricultural and Biological Engineers Vol. 49(1)* 75-82.

Flugstad, B. A. 2013. *Water Content and Electrical Conductivity Profile Measurements for Dispersive Media using Enhanced Time Domain and Frequency Domain Models.* Theisis : Oregon State Univerity, Department of Biological and Ecological Engineering .

Hamed, Y., P. Magnus, and R. Berndtsson. 2003. "Soil Solution Electrical Conductivity Measurements Using Different Dielectric Techniques." *SSSA J. Vol 67 No. 4* 1071-1078.

Hilhorst, M. A. 2000. "A Pore Water Conductivity Sensor." *Soil Sci. Soc. Am.* 64: 1922-1925.

Kahimba, F. C. ,R. Sri Ranjan* and M. Krishnapillai. 2007. "Impact of cable lengths on the accuracy of dielectric constant measurements by time domain reflectometry." *CANADIAN BIOSYSTEMS ENGINEERING* Vol. 49.

Logsdon, S. D. 2000. "Effect of Cable Length on Time Domain Reflectometry Calibration for High Surface Area Soils." *SSSA J. Vol. 64* 54-61.

Logsdon, S. D. 2005. "Time Domain Reflectometry Range of Accuracy for High Surface Area Soils." *Vadose Zone Journal 4:* 2004.

Pedregosa, et, al. 2011. "Scikit-learn: Machine Learning in Python." *Journal of Machine Learning Research 12* 2825-2830.

Regaladoa, C. M, R. Mun˜oz Carpenab, R. Socorroa, J.M. Herna´ndez Morenoc. 2003. "Time domain reflectometry models as a tool to understand the dielectric response of volcanic soils." *Geoderma 117* 313-330.

Robinson, D. A., S. B. Jones, J. M. Wraith, D. Or, and S. P. Friedman. 2003. "A Review of Advances in Dielectric and Electrical Conductivity Measurement in Soils Using Time Domain Reflectometry." *Vadose Zone Journal* 2: 444–475.

Topp, G.C., J.L. Davis, and A.P. Annan. 1980. "Electromagnetic deter- mination of soil water content: Measurement in coaxial transmis- sion lines." *Water Resour. Res. 16:* 574-582.

Topp, G.C., M. Yanuka, W.D. Zebchuk, and S.J. Zegelin. 1988. "Deter- mination of electrical conductivity using time domain reflectome- try: Soil and water experiments in coaxial lines." *Water Resour. Res. 24* (Water Resour. Res. 24) 945-952.

Topp, GC, J. L Davis, and A. P Annan. 1982. "Electricomagnetic Determination of Soil Water Content Using TDR: II. Evaluation of Instrallation and Configuration of Parallel Transmission Lines." *SSSA J. Vol 46* 678-684.