US Army Corps of Engineers®
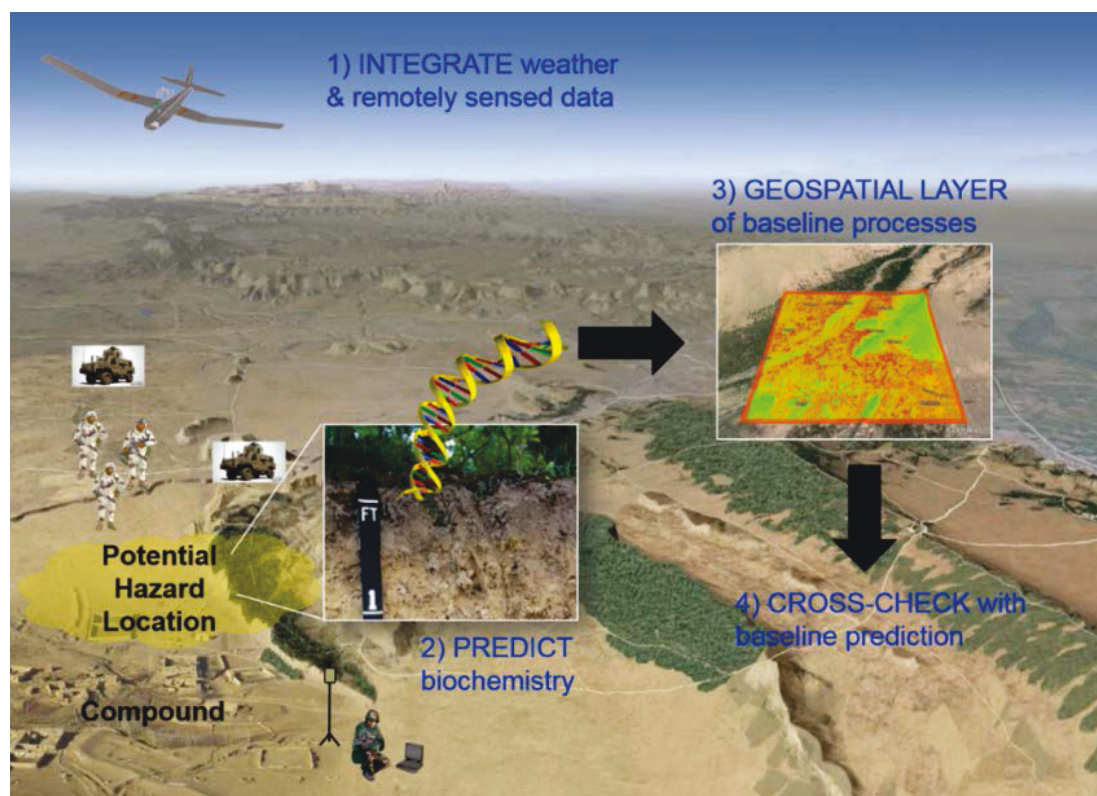Engineer Research and Development Center

ERDC 6.2 Geospatial Research and Engineering (GRE) ARTEMIS DRTSPORE

Dynamic Representation of Terrestrial Soil Predictions of Organisms' Response to the Environment

Robyn A. Barbato, Lauren Waldrop, Komi Messan, Robert Jones, Stacey J. Doherty, Karen Foley, Christopher Felt, Michael Morgan, and Youl Han

September 2018

Dynamic Representation of Terrestrial Soil Predictions of Organisms' Response to the Environment

Robyn A. Barbato, Lauren Waldrop, Komi Messan, Robert Jones, Stacey J. Doherty, Karen Foley, Christopher Felt, Michael Morgan, and Youl Han

Cold Regions Research and Engineering Laboratory
U.S. Army Engineer Research and Development Center
72 Lyme Road
Hanover, NH 03755

Final report

Abstract

The Dynamic Representation of Terrestrial Soil Predictions of Organisms' Response to the Environment (DRTSPORE) platform was created to *characterize* soil activities in barren, vegetated, and complex environments, *predict* biological-impacted processes in soil, and *generalize* the framework to solve a range of Army-relevant problems. Key biological processes include soil stabilization, emerging power sensors, brown out abatement, pathogen emergence, and contaminant and threat degradation. These capabilities are a new essential aid to national security, as they currently do not exist. The primary objective of this research effort, DRTSPORE, was to develop an environmental intelligence tool that adds a biochemical layer to current high resolution remotely sensed terrain and sophisticated weather model products. The mathematical models developed serve as libraries added to the existing graphic user interface. The improved environmental intelligence platform will provide Commanders a tactical decision aid to make better informed decisions about mobility, the placement and construction of a forward operating base, the placement of sensors, and the avoidance of areas where there is a potential for exposure to mobilized toxic materials in the soil.

Contents

Report Documentation Page

Figures and Tables

Figures

Tables

Preface

This study was conducted for the U.S. Army Corps of Engineers (USACE) 6.2 Geospatial Research and Engineering (GRE) Program under Army Terrestrial-Environmental Modeling and Intelligence System (ARTEMIS), "Dynamic Representation of Terrestrial Soil Predictions of Organisms' Response to the Environment" (DRTSPORE)". The technical monitor was Mr. John Eylander (CEERD-RR).

The work was performed by the Biogeochemical Sciences Branch (CEERD-RRN) and the Signature Physics Branch (CEERD-RRD) of the Research and Engineering Division (CEERD-RR), U.S. Army Engineer Research and Development Center, Cold Regions Research and Engineering Laboratory (ERDC-CRREL). At the time of publication, Dr. Justin Berman was Chief, CEERD-RRN; Dr. Andrew Niccolai was Chief, CEERD-RRD; and Mr. J. D. Horne was Chief, CEERD-RR. The Deputy Director of ERDC-CRREL was Mr. David B. Ringelberg, and the Director was Dr. Joseph L. Corriveau.

COL Ivan P. Beckman was the Commander of ERDC, and Dr. David W. Pittman was the Director.

Acronyms and Abbreviations

Ψ_m	Matric Potential
ATCC	American Type Culture Collection
Ca	Calcium
CEC	Cation Exchange Capacity
CO_2	Carbon Dioxide
CRREL	Cold Regions Research and Engineering Laboratory
Cu	Copper
ERDC	U.S. Army Engineer Research and Development Center
GRE	Geospatial Research and Engineering
K	Potassium
L	Loam
L-OM	Loam – Organic Matter
Mg	Magnesium
N	Nitrogen
ND	Not Detected
OM	Organic Matter
OTU	Operational Taxonomic Unit
P	Phosphorus
Q_{10}	Temperature Coefficient
qPCR	Quantitative Polymerase Chain Reaction
SL	Sandy Loam
SL+OM	Sandy Loam + Organic Matter
SL-OM	Sandy Loam – Organic Matter
TSP-SA	Temporal Predictions of Soil Activities
Zn	Zinc

Unit Conversion Factors

Multiply	By	To Obtain
British thermal units (International Table)	1,055.056	joules
cubic inches	1.6387064 E-05	cubic meters
degrees Fahrenheit	(F-32)/1.8	degrees Celsius
horsepower (550 foot-pounds force per second)	745.6999	watts
inches	0.0254	meters
ounces (mass)	0.02834952	kilograms
pounds (mass) per cubic foot	16.01846	kilograms per cubic meter
pounds (mass) per cubic inch	2.757990 E+04	kilograms per cubic meter
yards	0.9144	meters

1   Introduction

1.1   Background

Dynamic Representation of Terrestrial Soil Predictions of Organisms' Response to the Environment (DRTSPORE) is a comprehensive applied research work unit under the 6.2 Geospatial Research and Engineering Army Terrestrial-Environmental Modeling and Intelligence System (ARTEMIS) work package. Its purpose is to create a tactical decision aid that forecasts biological activity as a geospatial layer on a terrain map (Figure 1). It uses remotely sensed terrain data and weather models to predict soil function. Soil function ranges from general soil activity to specific biochemical function acquired through high resolution datasets. These datasets pertain to the transformation and fate of military relevant compounds of concern. Both the general and specific biochemical data were then mathematically modeled. The modeling frameworks within the platform ranged from a nonlinear approach, population-based model, or probabilistic simulated models in order to break through different modeling shortcomings and obtain robust results. The models are stored in the DRTSPORE Platform as libraries or look-up tables. Through the graphical user interface in ArcMap, the DRTSPORE platform acquires soil temperature and moisture data and then provides predictions of soil function. The DRTSPORE platform will provide Commanders a visualization platform to assess mobility, the placement and construction of a forward operating base which may be built by expeditionary forces, the placement of sensors, and the avoidance of areas where there is a potential for exposure to mobilized toxic materials in the soil.

DRTSPORE has six components listed below developed by the CRREL Soil Microbiology team.

1.   Testbed for Soil Science Research to acquire environmental measurements and test sensors
2.   Empirical Data Collection to develop a model of soil activity and create a robust dataset of microbial composition
3.   Mathematical Model of Soil Activity to predict soil activity and microbial trends according to soil temperature and moisture
4.   DRTSPORE Platform geospatial intelligence decision aid

5. Validation of Soil Activity using in situ field measurements of soil respiration

6. Sensor Technology — development and placement of in situ microbial fuel cells.

Figure 1. DRTSPORE Platform Concept to provide Army capability to inform persistence of hazards across the landscape and baseline soil activity.
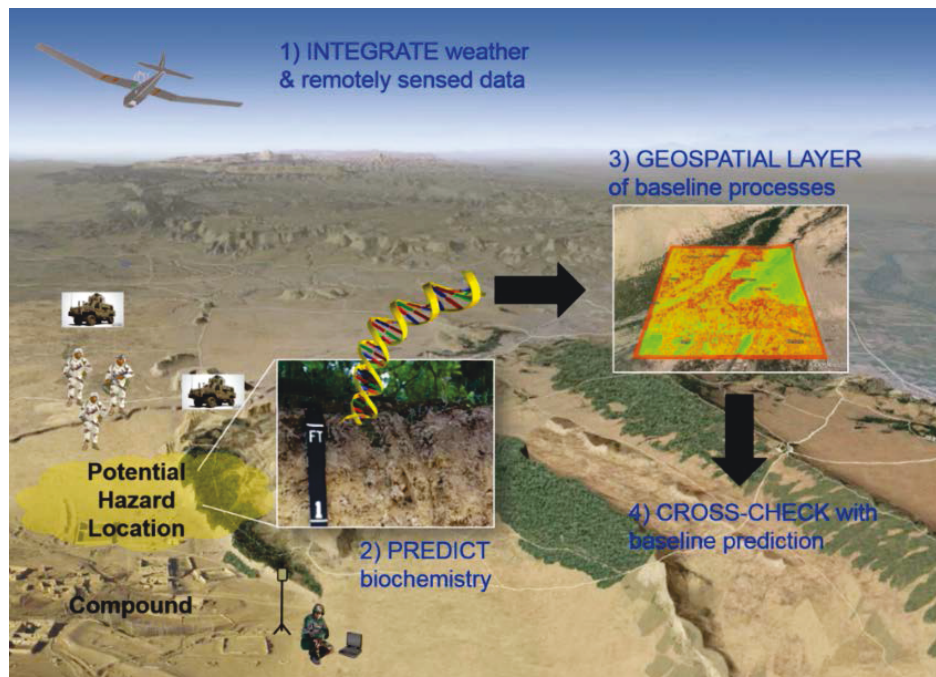
1.2   Objective

The ultimate goal of this research was to build an environmental intelligence tool that added a biochemical layer to current high resolution terrain and sophisticated weather model products.

1.3   Approach

This tool was designed to ingest weather model output from Terrestrial-Environment Rapidly-Relocatable Simulation (TERRASIM) within ARTEMIS. The weather model output is used to create soil temperature and moisture predictions forecasted out for three days through the GeoWATCH tool, developed by CREARE through a Small Business Innovation Research under ARTEMIS. The soil moisture data are robust at 30 cm resolution grid space and are validated by the Terrain Phenomenology and data Collection (TerraPAC) under ARTEMIS. This confirmation enables

better predictions of soil moisture and in turn soil activity. The soil temperature resolution is at the 1 km grid space, which is coarser and further resolution is required to provide more accurate predictions of soil activity. The final output of the platform is hosted in ArcMap in red-amber-green output to coordinate with the Signal Physics Representation in Uncertain and Complex Environments (SPRUCE) sensor placement within ARTEMIS. Additionally, the Soil Microbiology DNA Tracking Technology, to assess the utility of using microbes to track materials in lofted soil, serves as direct input to high resolution modeling products under the Dust-TERRASIM work unit. Finally, the plume modeling of biological and chemical hazards through SCIPUFF was implemented through the SPRUCE work unit.

1.4 Army relevance

The DRTSPORE platform was created to *characterize* soil activities in barren, vegetated, and complex environments, *predict* biological-impacted processes in soil, and *generalize* the framework to solve a range of Army-relevant problems. Key biological processes include soil stabilization, emerging power sensors, brown out abatement, pathogen emergence, and contaminant and threat degradation (Figure 2). These capabilities currently does not exist and are a new essential aid to national security. The CRREL DRTSPORE System is the first of its kind mainly because of the technical challenges associated with modeling complex and dynamic biological processes occurring at the soil surface.

Figure 2. Army relevant biological processes relevant to the ARTEMIS intelligence system.
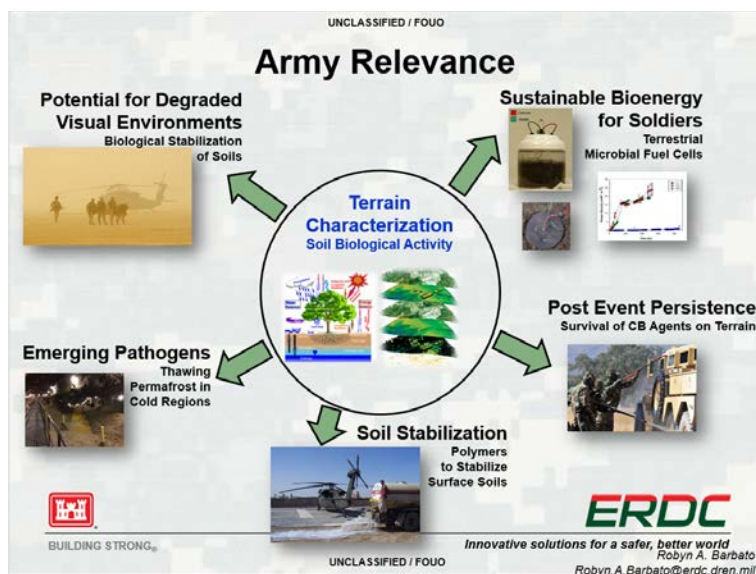
2 CRREL Soil Microbiology Testbed for Soil Science Research

In 2011, the ERDC CRREL Soil Microbiology Field Site was created within the CRREL campus in Hanover, NH. Former uses at the site include countermine and night vision activities. Our goal was to prepare the field site so that multiple soils with varying characteristics could be tested under one weather regime. In phase one, six soil lanes were prepared (Figure 3). During late 2011, three lanes 2.4 m wide and 18.3 m long were cleared of sod and the existing soil smoothed (Figure 3A). At the northern end of the lanes, a trench was dug to provide the effects of slope for potential studies. Lanes were separated by a 1.5 m wide undisturbed strip. New experiments required studying applied materials and microbial response in different soil types and how the physical properties of those soil types impact the system. In spring 2012, three more soil lanes were constructed using imported soils. To create each imported soil lane a 0.6 m deep trench, 2.4 m wide by 18.3 m long, was excavated The imported soils were placed into the respective lanes, with mechanical compaction between loads, until the lanes were level with the surrounding landscape. The lanes are separated by 2.4 m wide undisturbed strips. Temperature and moisture probes (Onset HOBO probes, Bourne, MA, USA) were installed in these lanes, autumn of 2016.

Figure 3. Preparation of the field site. A) Heavy equipment used to create soil lanes. B) Hole for soil moisture and temperature probes. C) Data loggers for probes. D) Installation of towers to host sensor-suites. E) Field site with sensors.

Soil temperature and moisture probes were installed at 16 locations within and between the lanes to provide data for both vegetated and cleared soil (Figure 3B, C). HOBO U30 data loggers with associated Onset 12-bit temperature smart sensors and soil moisture smart sensors were installed at depths of 2.5 cm, 5 cm, 10 cm, 15 cm and 20 cm. Using an auger, holes were excavated for installation of the probes (Figure 3B). The soil surrounding each probe was sieved when installing and the soil between probe depths was not sieved. All soil used in probe installation was of the lane's soil type. After all probes for a location were installed, a shallow trench was dug from the probe site to a metal pole, upon which a datalogger was mounted (Figure 3C). For future permanent mounting of data collection devices, such as terrestrial laser scanners, instrument towers were installed at two locations within the site in autumn of 2012 (Figure 3D). To ensure a stable base for accurate data collection, each tower was embedded in a 1.5 m deep concrete footing.

Maintenance strategies to keep the soil lanes barren include weeding the lanes, applying glyphosate if necessary, and mowing the areas between the lanes.

2.1 Water release curves for each soil

In the laboratory, a water release curve was obtained for the SL+OM, SL, L-OM, and SL-OM soils. To do so, soils were saturated and then held at a particular pressure, or matric potential using a pressure kettle device. The gravimetric water content was measured at each point. Four matric potential values were tested for each soil to develop the water release curve. With this information, the water content of the soil could be transformed to matric potential, which is a more meaningful measure of soil water that is available to organisms (Moyano et al. 2013). As seen in Figure 4, each soil has its own water release curve. At a matric potential of -33 kPa, the optimal water availability for organisms, it is important to note that each soil has a very different gravimetric water content. The SL+OM soil has the best water holding capacity, with the SL soil slightly behind (Figure 4). The low organic matter soils, L-OM and SL-OM, have a reduced ability to hold water (Figure 4).

Figure 4. Water release curves for determining matric potential of
L-OM, SL, SL+OM, and SL-OM.

L-OM = $1.5311x^{-1.909}$
$R^2 = 0.9808$

SL = $4.5651x^{-1.957}$
$R^2 = 0.706$

SL+OM = $0.8596x^{-2.95}$
$R^2 = 0.769$

SL-OM = $0.3822x^{-2.512}$
$R^2 = 0.8981$
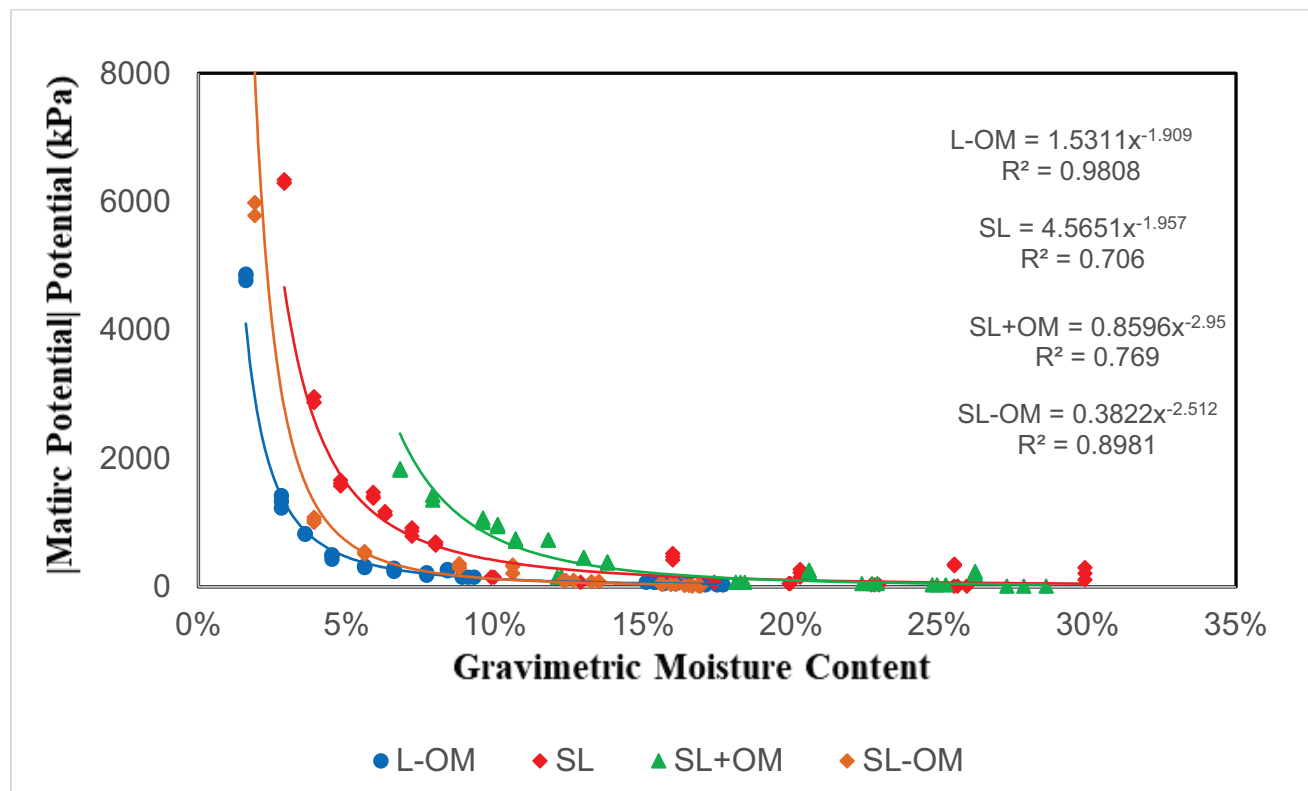
2.2 Soil temperature and moisture data collection

Since its inception in 2011, soil moisture and temperature data points have
been collected every 15 minutes at depths of 2.5 cm, 5 cm, 10 cm, 15 cm
and 20 cm below the soils' surface. It should be noted that while the soil
temperature probes accurately measure temperatures below freezing, the
soil moisture probes are only accurate when soil water is in liquid form.

Here, we present soil temperature and moisture data from 2013 to 2017 at
depths of 2.5 cm, 5 cm, and 10 cm (Figure 5 to Figure 10). As shown in the
temperature graphs (Figure 5 to Figure 7), the spring thaw varied from
year to year. In 2017, when all of the soils were measured, the trend per-
sisted regardless of soil type. From the graphs it can be seen that across
the four soil lanes, the temperature remains fairly consistent between all
four soil types. The greatest diurnal temperature fluctuation occurred at
the soil surface during the warmer months of each year. Soil moisture,
however, varied across the soil lanes due to the differing water holding ca-
pacities of each soil type (Figure 8 to Figure 10). The water holding capac-
ity of a soil is dependent primarily upon soil texture and organic matter

content. When a soil has a high sand content and low silt, clay, and organic matter, it will have a low water holding capacity compared to soils that are able to maintain saturation longer with higher clay, silt and organic matter content.

Figure 5. Soil temperature data from all four soil types at a depth of 2.5 cm.
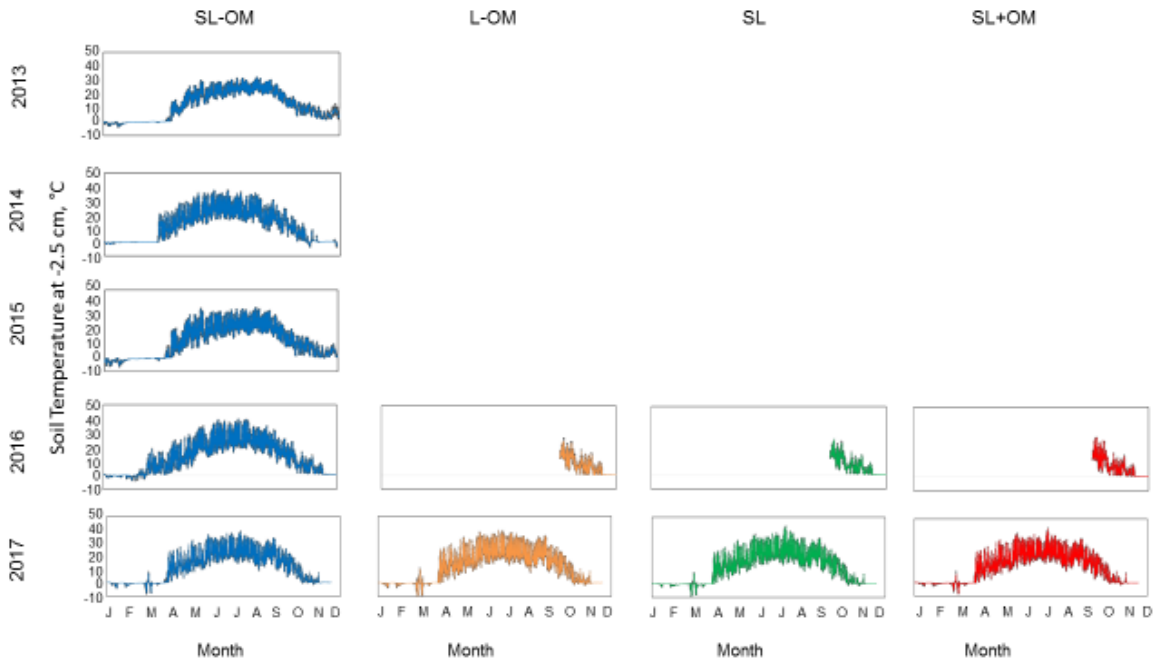
Figure 6. Soil temperature data from all four soil types at a depth of 5 cm.
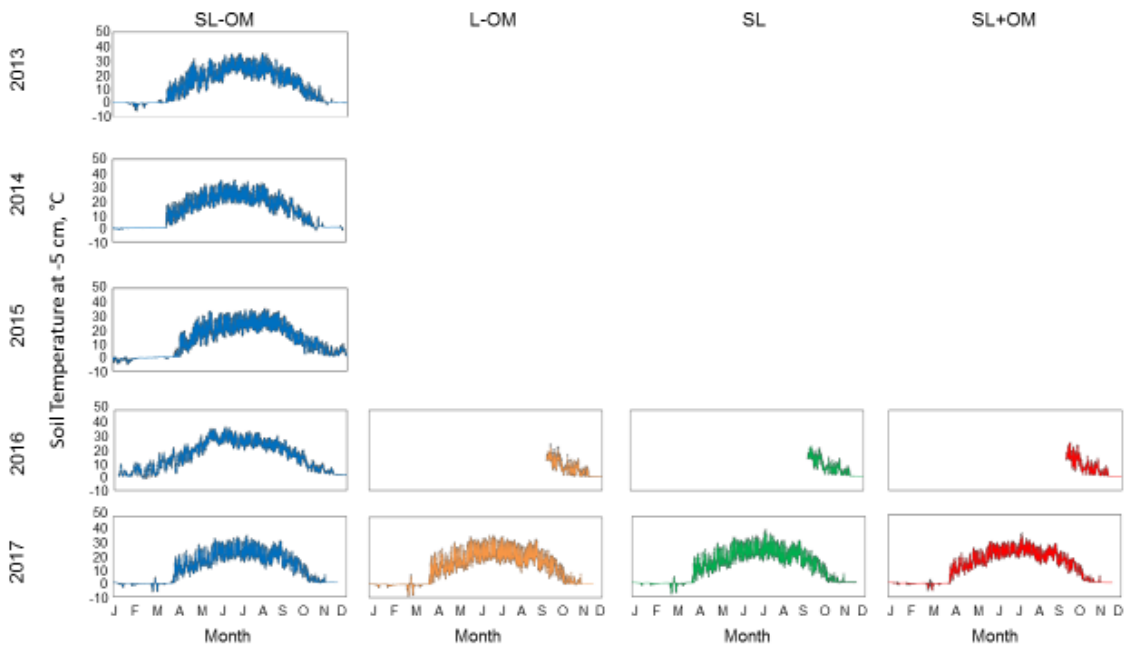
Figure 7. Soil temperature data from all four soil types at a depth of 10 cm.
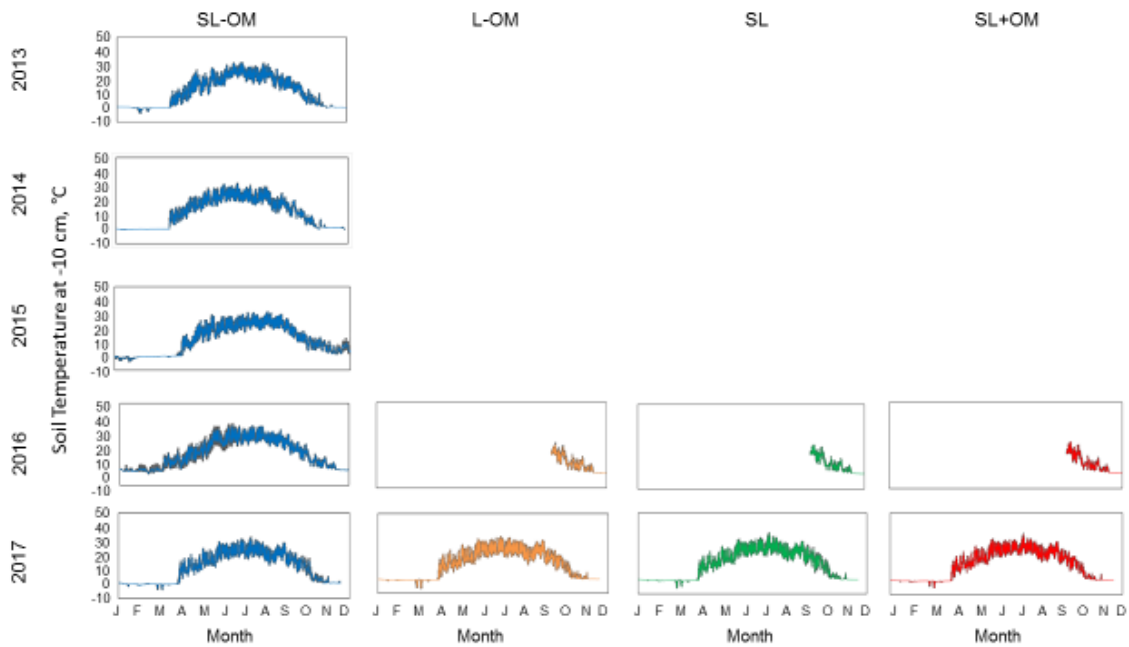
Figure 8. Soil moisture data from all four soil types at a depth of 2.5 cm.
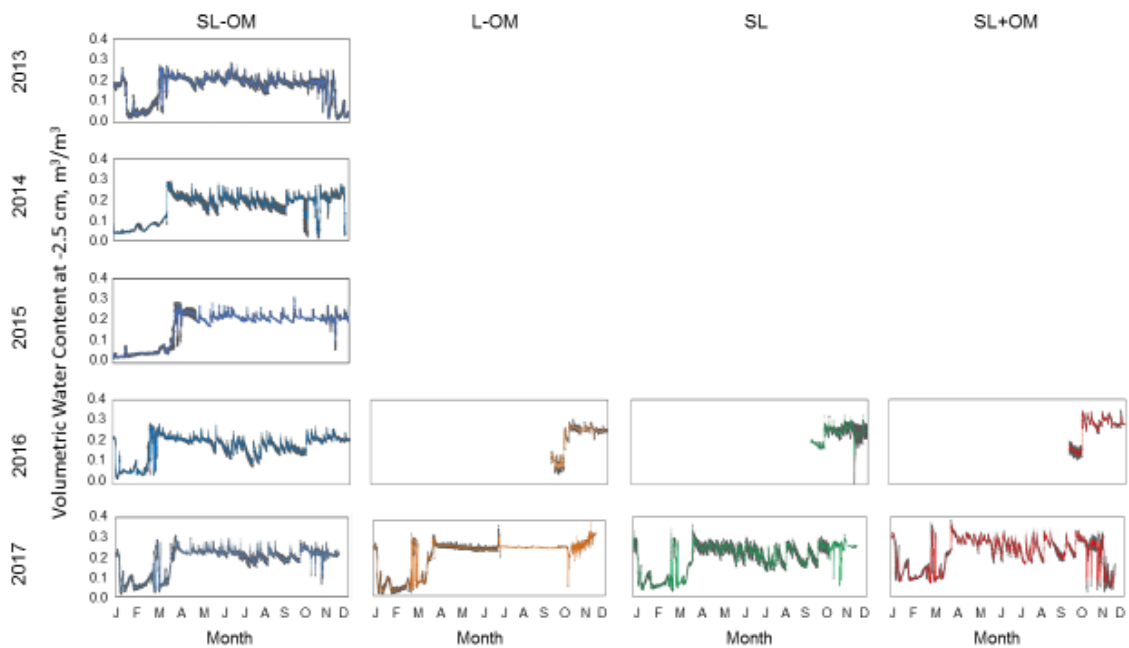
Figure 9. Soil moisture data from all four soil types at a depth of 5 cm.
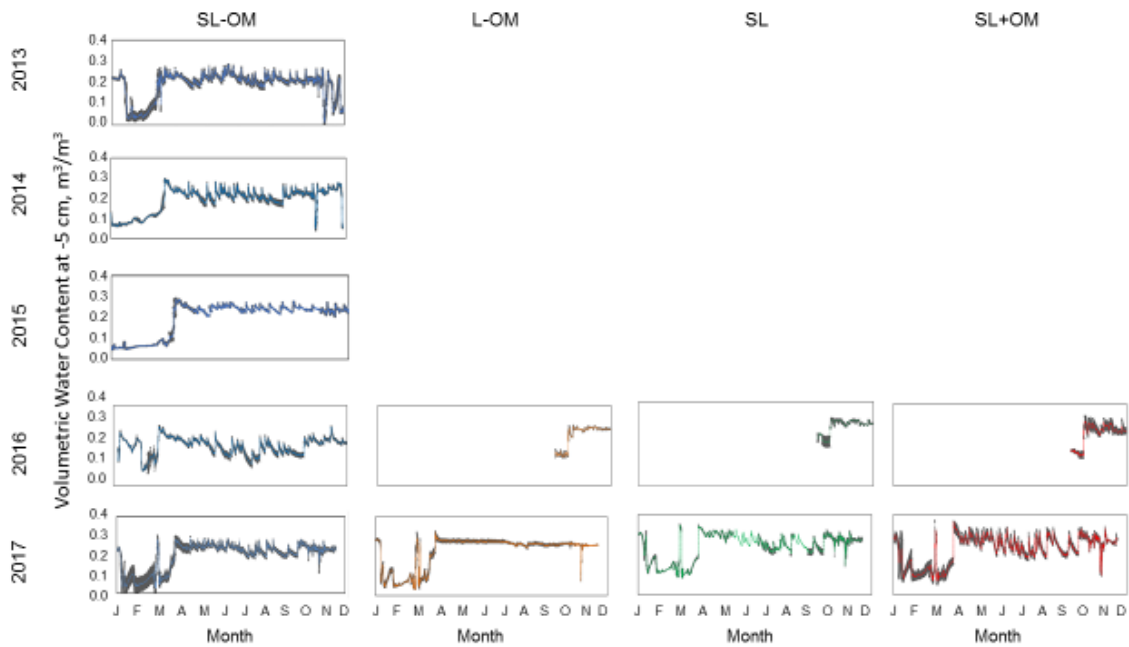
Figure 10. Soil moisture data from all four soil types at a depth of 10 cm.
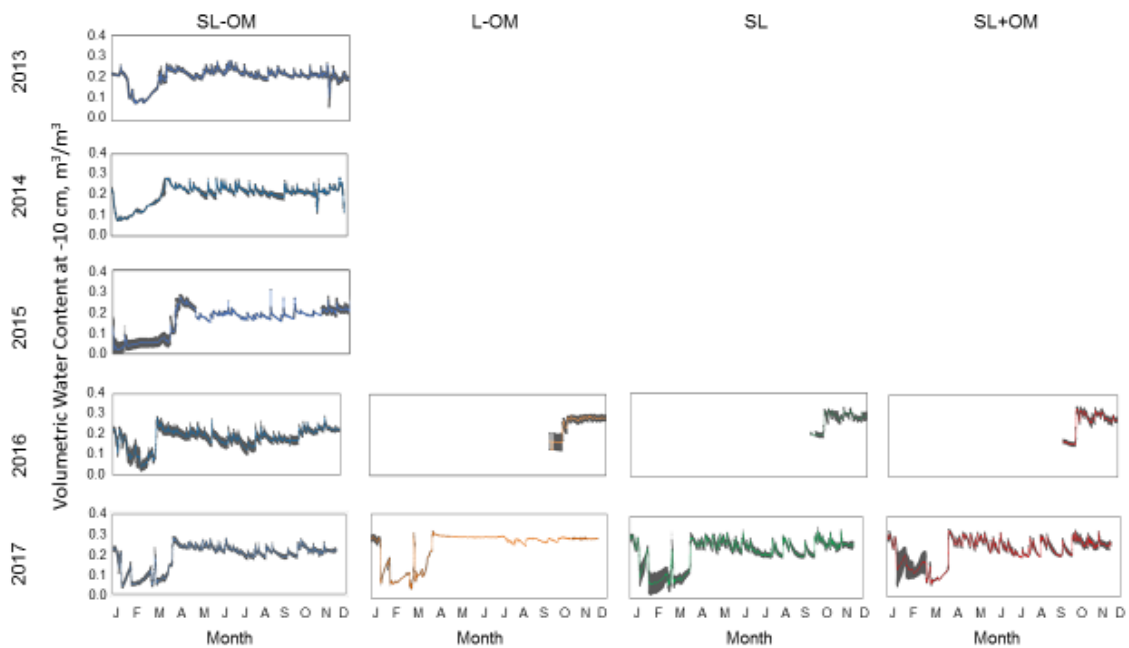
3 CRREL Soil Microbiology Empirical Data Collection

A laboratory incubation study was designed to collect soil biological attributes at varying conditions of soil temperature, moisture, and organic matter content, which are critical drivers of microbial community composition. Soil biological attributes measured include soil respiration, through the efflux of CO_2, as well as bacterial and fungal microbial communities through DNA sequencing. Only bacteria will be discussed in this section. Detailed information of the experimental set-up and CO_2 data can be found in Barbato et al. (2015).

The aim of this task was to use DNA technology to assess changes in bacterial taxonomy according to the changing variables of soil moisture and temperature. High resolution datasets such as the bacterial DNA dataset could be added to the DRTSPORE Platform. The purpose of this study was to investigate if changes in microbial communities were predictable and consistent. To do so we combined high throughput DNA sequencing techniques with bioinformatics and data analysis to visualize the communities from multiple soil types across the temperature and moisture gradient. We expected that alterations in temperature and moisture would have pronounced effects on the composition of the soil microbial community. It is possible that these alterations would be so pronounced they would create unique bacterial communities that could serve as a fingerprint for their environmental condition of origin. We explored the both the validity and composition of these potential fingerprint communities through supervised learning methods (algorithms that use known data to make predictions about unclassified data) and taxonomic classification through high throughput sequence alignment analysis. Together, these methods would identify which environment condition, if any, had the strongest impact on microbial diversity i.e., created the most unique community and which microbial taxa made up that community.

3.1 Laboratory incubation set-up

Samples for this study originated from four soils at the CRREL Soil Microbiology Testbed. They include a sandy loam with an organic matter amendment (SL+OM), an unamended sandy loam (SL), a loam with low detectible concentrations of organic matter (L-OM), and a sandy loam with very little to no detectible organic matter (SL-OM). For each soil

group, four temperatures and five moisture conditions were tested: 5°C, 15 °C, 25 °C, and 30 °C and -500, -100, -33, -10, and -5 kPa, respectively. Samples were incubated in quadruplicate while connected to a Micro-Oxymax Respirometer (Columbus Instruments, Columbus OH, USA) to measure CO_2 efflux. At the end of the incubation, they were destructively sampled for DNA extraction (Barbato et al. 2015).

3.2 DNA extraction and sequencing

Genomic DNA was extracted from soil using the PowerSoil DNA Isolation Kit and protocol (MoBio Laboratories, Carlsbad, CA). DNA concentration was measured using a Qubit 3.0 Fluorometer (ThermoFisher Scientific, Grand Island, NY) and adjusted to standardize the concentration within a treatment set. Sequence amplification was performed to target the 16S rRNA gene for bacteria. Sequences were run on a MiSeq Sequencer (Illumina, San Diego, CA) at Argonne National Laboratory using the paired end protocol, producing forward and reverse readings.

3.3 Taxonomic and diversity analysis methods

Sequences were processed through a pipeline within the open source program Quantitative Insights Into Microbial Ecology (QIIME) (Caporaso et al. 2012). The forward and reverse sequences were mapped to overlapping sections and joined together into one sequence using fastq-join (Aronesty 2011). The joined reads were then quality filtered and demultiplexed to match their original samples. Using the QIIME open reference picking workflow, taxonomy in the form of operational taxonomic units (OTUs) was assigned at 97% identity, the database used for taxonomic assignment was the Greengenes database (DeSantis et al. 2006). OTUs are a taxonomic clustering classification scheme that allows us to make probabilistic assumptions on the identity of specific sequences and group highly similar sequences together to a common identity.

Beta Diversity metrics, how similar or disparate samples within and between treatments, were analyzed using the core diversity pipeline script in QIIME, resulting in a summary table of the assigned taxa and their relative abundance proportional to total OTUs assigned per sample. Principle coordinate analysis plots and coordinate files were calculated using the weighted unifrac metric for bacteria (Lozupone et al. 2005).

3.4   Supervised learning methods

We tested the potential for soil microbial communities to serve as indicators of environmental conditions using supervised learning methods. In our attempts to use machine learning to classify the samples, we used the random forests (Breiman 2001) algorithm which utilizes hundreds of decision trees grown on random subsampling of OTUs from each sample to understand which mixture of OTUs were unique to each classification. OTU tables from samples were separated into their respective soil groups (SL+OM, SL, L-OM, and SL-OM) and the temperature and moisture variables were tested individually. For each of the four temperatures and five moistures within each soil set, the random forest algorithm generated 500 decision trees, each tree being a random subsampling of OTUs from the sample. For each unit, the algorithm judged the importance of that OTU for accurate classification of the sample (i.e., if you remove 'species X' from this decision tree do you still arrive at an accurate classification?) After 500 iterations, the algorithm produced a binary decision tree diagram which was then used as a model to classify the original data. The validation of the classifier was checked via ten fold cross validation where the original data were randomized and divided into a training set and a test set. The training set was used to generate the model via random forests. Once the model was complete it was validated against the test data set for accuracy and repeated ten times. Afterwards a confusion matrix (a comparison of the actual vs predicted classification) was generated to demonstrate the predictions and accuracy of the model.

3.5   Class level significant taxonomic analysis methods

After receiving the results of the supervised learning, we assessed the patterns of the microbial taxa using temperature as the basis for statistical comparison. Summarized relative abundance OTU tables of each soil at the class level were filtered to only include units with positive identification (i.e., hits with no distinctive class or defined as "other" were removed) and recalculated to equal 100%. Using these identifications an analysis of variance was performed using the microbial abundance, as the y variable, and temperature, as the x variable. Based on the probability of significance from the variance analysis, the top 15 most abundant taxa with a probability of less than 0.05 were selected and normalized to 100% for each soil class. They were then grouped by phylum and plotted as stacked bar charts.

3.6 Supervised learning results: Are the communities according to temperature and moisture unique enough to allow for accurate classification?

We expected the random forest algorithm to be able to determine the mixture of OTUs unique to a temperature or moisture group in order to classify them accordingly. The temperature and moisture confusion matrices identify the accuracy of random forest model by taking the ratio of the correct predictions to the incorrect classifications, with 100% indicating that the model correctly identified each sample according to its environmental variable without knowing the identity.

For SL+OM, classification according to temperature was largely successful having an average accuracy of 85% (Table 1). The errors in the predictions likely resulted from the classifier mistaking a sample for one at a slightly lower temperature. For SL+OM, the classifier had a much more difficult time with moisture and was only able to achieve an average accuracy of 72% (Table 1). Only the low moisture condition was predicted with 100% accuracy indicating a distinctive community (Table 1).

Table 1. Random forest classification confusion matrices for SL+OM.
The header row indicates the actual identity of the sample whereas the first column indicates the predicted identity. The color shading indicates the weight of the count (hotter indicating more counts). Accuracy corresponds to the ratio of correct classifications to total classifications.

Temperature	05°C	15°C	25°C	30°C	Accuracy (%)
05°C	19	0	1	0	95
15°C	6	14	0	0	70
25°C	1	0	19	0	95
30°C	0	0	4	16	80

Moisture	1 Very Low	2 Low	3 Medium	4 High	5 Very High	Accuracy (%)
1 Very Low	10	3	0	0	3	62.5
2 Low	0	16	0	0	0	100
3 Medium	0	1	8	4	3	50
4 High	0	0	2	12	2	75
5 Very High	3	0	1	0	12	75

For SL, the average accuracy for temperature predictions was low (78%), but the classifier accurately identified the 30 °C samples (Table 2). However, misclassifications do not seem to follow a pattern for the other temperatures. For moisture, the classifier performed poorly with an average of 44% accuracy, indicating that moisture changes did not promote disparate communities in the SL soil in a repeatable way (Table 2).

Table 2. Random forest classification confusion matrices for SL.
The header row indicates the actual identity of the sample whereas the first column
indicates the predicted identity. The color shading indicates the weight of the count
(hotter indicating more counts). Accuracy corresponds to the ratio of correct
classifications to total classifications.

Temperature	05°C	15°C	25°C	30°C	Accuracy (%)
05°C	15	4	1	0	75
15°C	0	16	4	0	80
25°C	1	2	13	4	65
30°C	0	0	0	20	100

Moisture	1 Very Low	2 Low	3 Medium	4 High	5 Very High	Accuracy (%)
1 Very Low	5	3	0	3	5	31.25
2 Low	1	10	1	1	3	62.5
3 Medium	3	4	2	5	2	12.5
4 High	2	1	7	5	1	31.25
5 Very High	2	0	1	0	13	81.25

The classifier had the greatest success with the L-OM soil with the exception of the 25 °C samples (Table 3). For the 25 °C samples, there may have been many microbial communities that were abundant, so no particular group stood out as a source of classification. Contrary to the moderate success with temperature, the moisture based classifier on average had more misclassifications than correct classifications compounding evidence that moisture is not as strong a variable for classification (Table 3).

Table 3. Random Forest Classification Confusion Matrices for SL.
The header row indicates the actual identity of the sample whereas the first column
indicates the predicted identity. The color shading indicates the weight of the count
(hotter indicating more counts). Accuracy corresponds to the ratio of correct
classifications to total classifications.

Temperature	05°C	15°C	25°C	30°C	Accuracy (%)
05°C	20	0	0	0	100
15°C	2	18	0	0	90
25°C	8	4	8	0	40
30°C	0	0	2	18	90

Moisture	1 Very Low	2 Low	3 Medium	4 High	5 Very High	Accuracy (%)
1 Very Low	10	3	3	0	0	62.5
2 Low	2	10	4	0	0	62.5
3 Medium	0	7	8	1	0	50
4 High	0	4	3	6	3	37.5
5 Very High	0	1	6	1	8	50

The SL-OM soil had the lowest average accuracy (75%) for any of the temperature-based classifications (Table 4). However, all misclassifications occurred within one step (higher or lower) of the actual temperature. This could be indicative of overlapping communities between the temperature regimes. Similar to the other soils tested, moisture was difficult to classify in the SL-OM soil, with an average accuracy of 48% (Table 4).

Table 4. Random forest classification confusion matrices for SL-OM.
The header row indicates the actual identity of the sample whereas the first column
indicates the predicted identity. The color shading indicates the weight of the count
(hotter indicating more counts). Accuracy corresponds to the ratio of correct
classifications to total classifications.

Temperature	05°C	15°C	25°C	30°C	Accuracy (%)
05°C	15	5	0	0	75
15°C	0	14	6	0	70
25°C	0	0	17	3	85
30°C	0	0	6	14	70

Moisture	1 Very Low	2 Low	3 Medium	4 High	5 Very High	Accuracy (%)
1 Very Low	11	5	0	0	0	68.75
2 Low	3	12	1	0	0	75
3 Medium	1	2	6	4	3	37.5
4 High	0	0	5	6	5	37.5
5 Very High	4	0	0	8	4	25

3.7 Class level significant taxa results: What are the significant taxa contributing to difference in classification?

Soil is one of the most diverse media in terms of microbial diversity, with upwards of 400 individual genera within a soil sample. However, while many members may be present, it is not necessarily indicative of their importance or contribution to the community function. A greater abundance of a microbial group usually indicates that is well situated to survive under the environmental conditions for that sample. As was indicated in the prior section, the community makeup may be unique to each environmental condition and could potentially be used to guide classification algorithms.

Analysis of Variance was used to obtain the top 15 significant bacterial taxa at the 95% confidence interval under the premise that they contributed to the majority of the total community. In these soils, certain phyla dominated the microbial profile for the entire sample regardless of temperature or moisture, which could be considered to be the core microbiome of that soil. Other microbial members only became abundant under certain conditions e.g., changes in temperature or moisture, particularly in the L-OM soil. Furthermore, small changes in temperature and moisture elicited a dramatic response in the microbial community.

For the SL+OM soil, Actinobacteria (orange) made up almost the entirety of middle three moistures in the 30 °C sample set (Figure 11). While Bacteroidetes was also abundant (10-30%), members within that group seemed to have moisture and temperature constraints above 5 °C (Figure

11). Acidobacteria, Gemmatimonadetes, and Planctomycetes were generally stable between temperature and moisture groups with the exception of the 30 °C samples which showed more moisture-dependent shifts. Soils incubating at 5 °C had the most stability in terms of similar community profiles across the moisture gradient (Figure 11).

For the SL soil, Acidobacteria and Proteobacteria were the principle players occupying between 20-60% and 20-30% of the community in each sample, respectively (Figure 12). With the exception of Actinobacteria and Chloroflexi, the other groups retained approximately the same proportions across temperature and moisture regimes only having individual shifts of 1-5% (Figure 12).

Similarly to SL+OM, the L-OM soil was largely dominated by Actinobacteria followed closely by Proteobacteria (Figure 13). The intergroup variability of Proteobacteria was far larger than Actinobacteria, as Beta-, Delta-, and Gamma-proteobacteria varied greatly with moisture and temperature (Figure 13). Actinobacteria was represented entirely by one member. Gemmatimonadetes was also a principle taxa particularly in the warmer samples while Chloroflexi become abundant in the middle temperatures (15-25°C, Figure 13).

Proteobacteria were most abundant in the SL-OM soil, which was similar to what occurred in the L-OM soil (Figure 14). Both of these soils have low organic matter content. However, unlike the L-OM, the Proteobacteria were the uncontested dominant group with the exception of the lower moisture tested (Figure 14). Under those conditions, Actinobacteria rose as high as 50% of the total community (Figure 14). Bacteroidetes showed a clear preference under higher moistures conditions and 15 °C (Figure 14). Examining Proteobacteria further we observed that the proportions of Alpha-, Beta-, and Delta-proteobacteria remained relatively stable (varying within 5-7%) across temperatures and moistures only, with changes in their patterns evident for the lowest moistures for the 25 °C and 30 °C samples (Figure 14).

Figure 11. Top 15 Significant Taxa for SL+OM. The size of each bar represents the percent abundance of that class and each bar is colored by phylum and shaded by class. Samples are divided into replicate (Ex: R1, R2), moisture (Ex: 1VL, 2L), and temperature (Ex: 15 °C, 25 °C). Moisture designations have been abbreviated for visualization purposes, 1VL (Very Low) corresponds to the driest moisture content whereas 5VH (Very High) corresponds to the wettest moisture content, all other moisture designations are in order from driest to wettest from left to right.
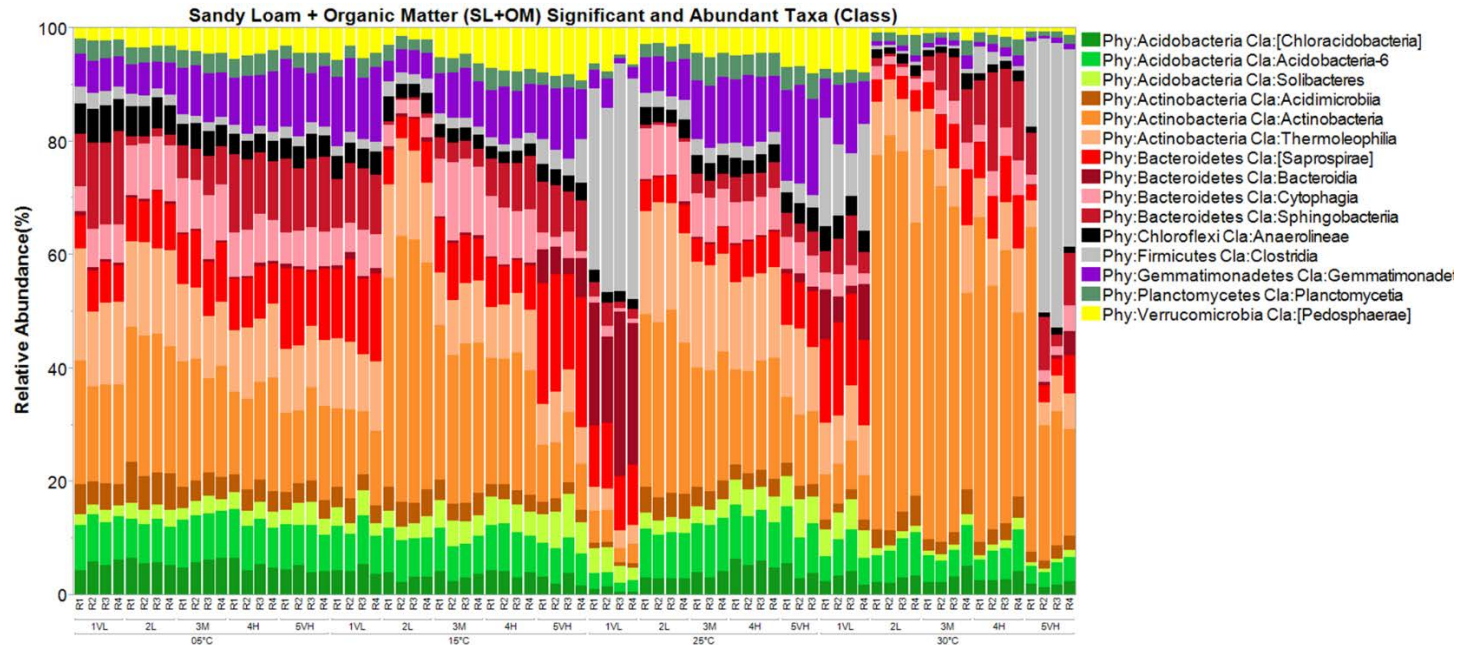
Figure 12. Top 15 Significant Taxa for SL. The size of each bar represents the percent abundance of that class and each bar is colored by phylum and shaded by class. Samples are divided into replicate (Ex: R1, R2), moisture (Ex: 1VL, 2L), and temperature (Ex: 15 , 25 °C). Moisture designations have been abbreviated for visualization purposes, 1VL (Very Low) corresponds to the driest moisture content whereas 5VH (Very High) corresponds to the wettest moisture content, all other moisture designations are in order from driest to wettest from left to right.
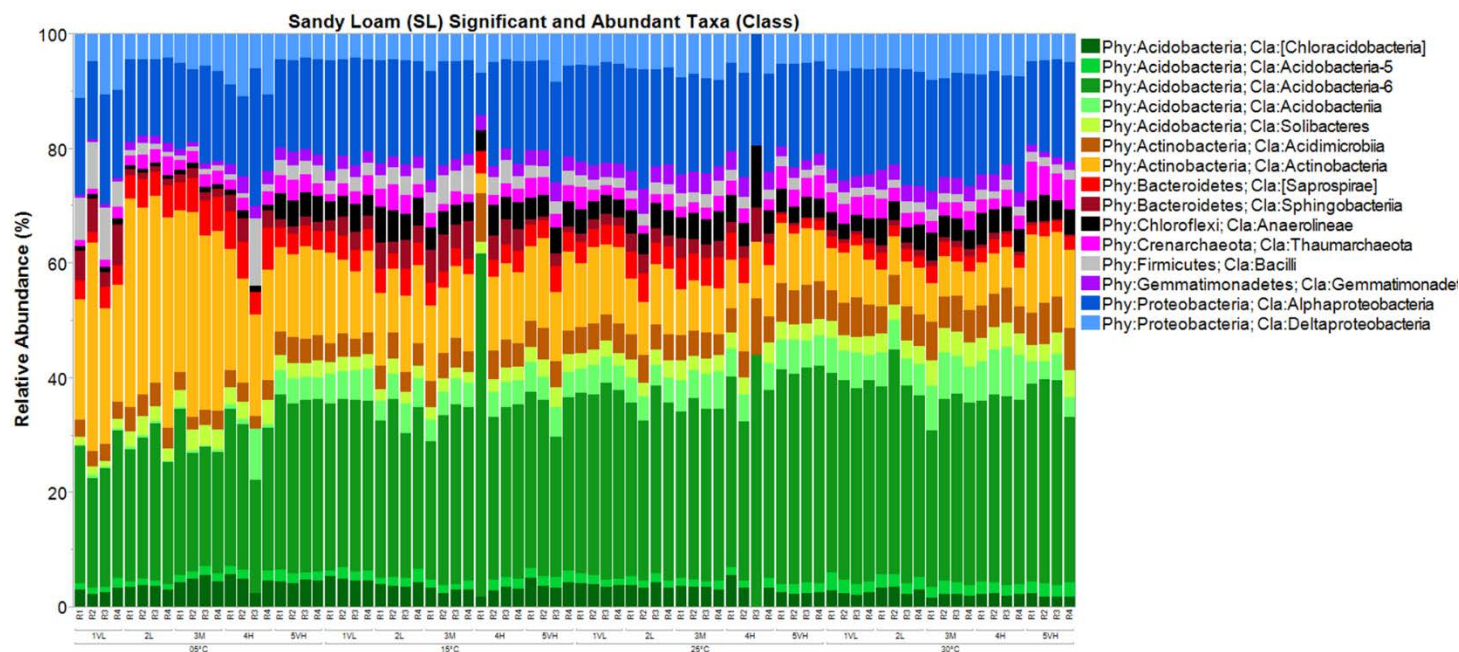
Figure 13. Top 15 Significant Taxa for L-OM. The size of each bar represents the percent abundance of that class and each bar is colored by phylum and shaded by class. Samples are divided into replicate (Ex: R1, R2), moisture (Ex: 1VL, 2L), and temperature (Ex: 15 °C, 25 °C). Moisture designations have been abbreviated for visualization purposes, 1VL (Very Low) corresponds to the driest moisture content whereas 5VH (Very High) corresponds to the wettest moisture content, all other moisture designations are in order from driest to wettest from left to right.
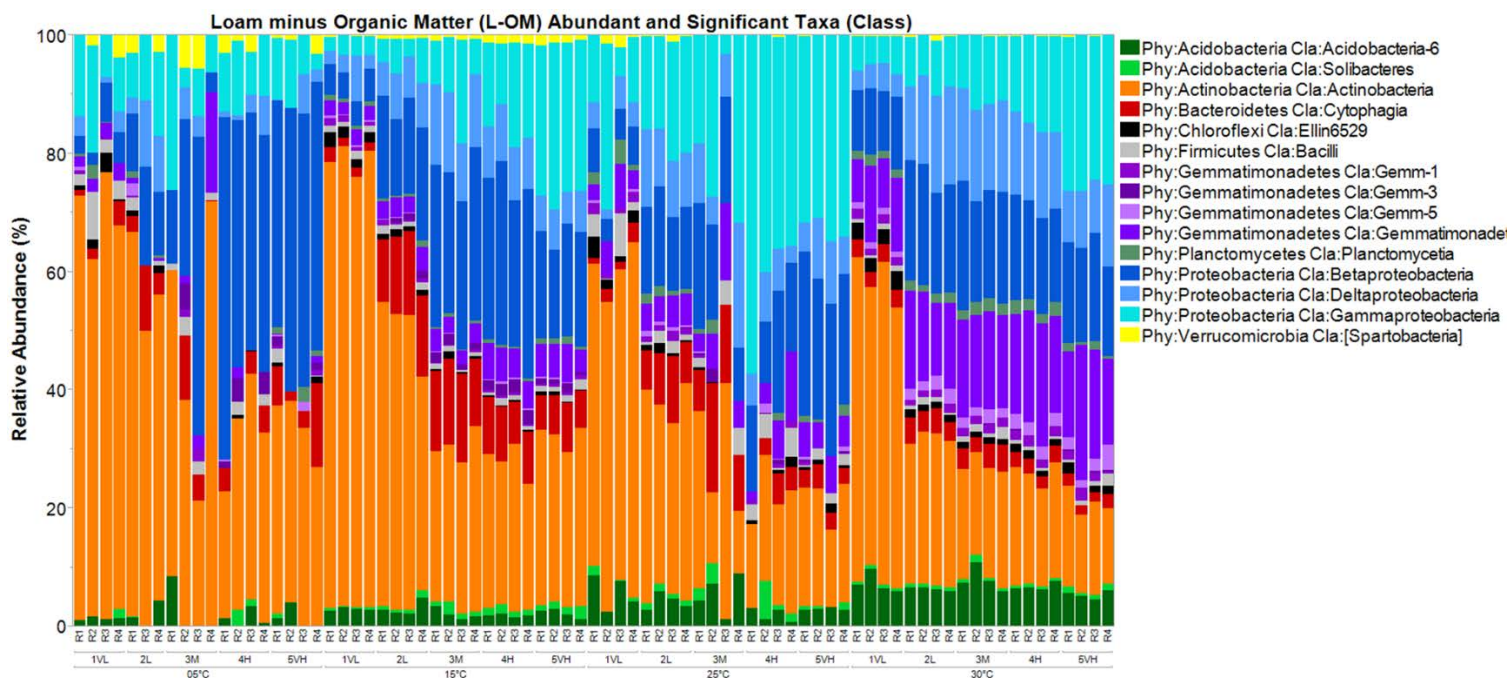
Figure 14. Top 15 Significant Taxa (SL-OM): The size of each bar represents the percent abundance of that class and each bar is colored by phylum and shaded by class. Samples are divided into replicate (Ex: R1, R2), moisture (Ex: 1VL, 2L), and temperature (Ex: 15 °C, 25 °C). Moisture designations have been abbreviated for visualization purposes, 1VL (Very Low) corresponds to the driest moisture content whereas 5VH (Very High) corresponds to the wettest moisture content, all other moisture designations are in order from driest to wettest from left to right.
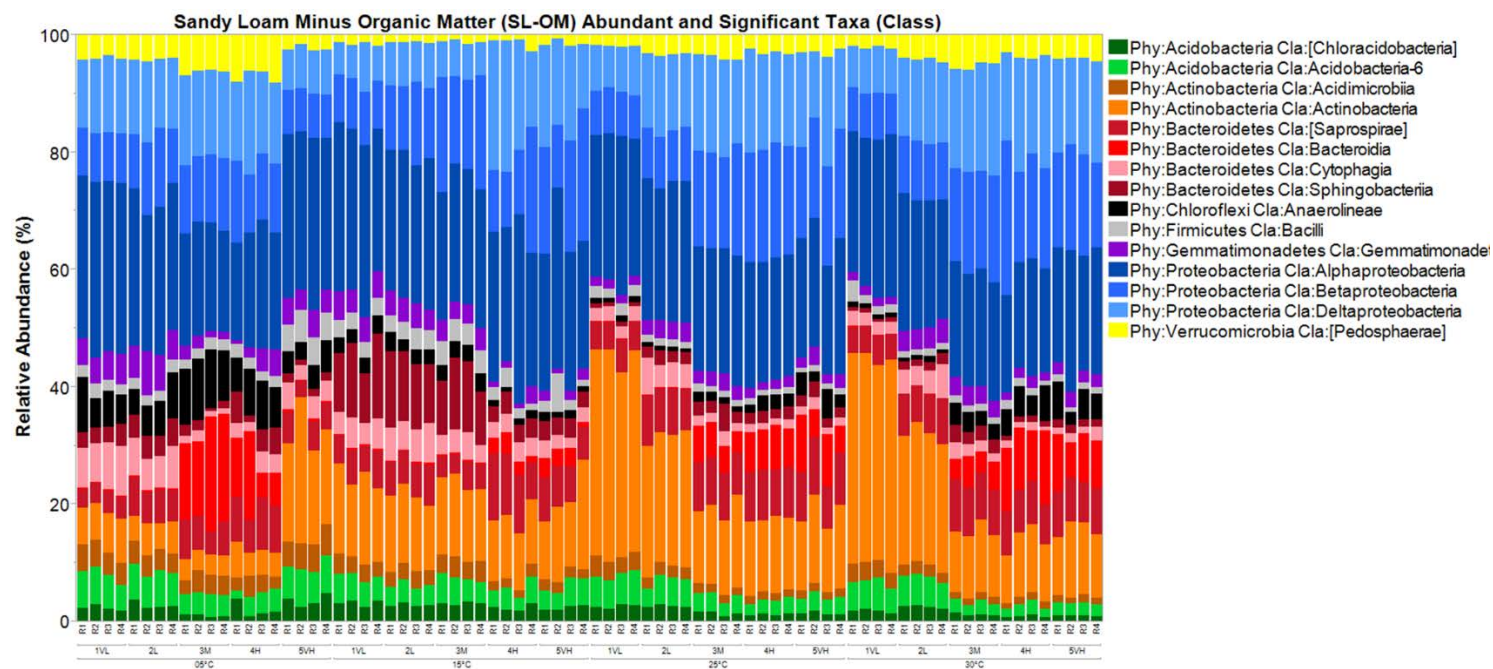
3.8   Discussion

Supervised learning algorithms are a powerful tool in classification of eco-
logical data to uncover the specific aspects of a sample that are defining or
unique. We aimed to determine which of the two environmental variables,
temperature and moisture, created more distinctive changes in microbial
community structure and the influence of soil type. The goal of the ran-
dom forest classification tests was to gain validation of the predictive
power of the classification of microbial communities. Variables that do not
elicit distinctive responses from the community resulted in misclassifica-
tion. The strength of temperature and moisture in regards to classification
was tested. For all soil the soil groups in this study, the average accuracy
was greatest in the temperature based classification, with the highest aver-
age being 85%. While around 90-95% would be best, in our study, changes
in microbial communities due to temperature were more effective indica-
tors than moisture. The accuracy was not ubiquitous for all temperatures,
5 °C and 30 °C for example had a consistently high accuracy across soil
types. This may be due to a few members that defined the community. In
addition to higher accuracy, temperature also has less predictive sway than
moisture. As was mentioned before, the misclassifications tended to fall on
temperatures within 1 step of the actual, e.g. classified as 5 °C or 25 °C
when in reality it was 15 °C, which could mean that with changing temper-
atures the communities were more defined but there was significant over-
lap. In general, moisture had a greater amount of predictive sway which
could be attributed to the biological importance of available moisture,
meaning that community members normally unable to survive at certain
temperatures could do so under optimal moisture conditions. To further
expand on the impact of temperature on the microbial community, we ex-
plored patterns in the class level taxa of the soil samples.

The profiles of significant taxa strongly varied across soil types. This kind
of feedback, knowing who the principal taxa are and how and why they
may change, is integral to the DRTSPORE Platform as it can inform the
state of the microbial community and its ability to respond to changes.
Each soil had one to three dominate groups making up 20-60% of the total
community that were consistent across the temperature and moisture re-
gimes, suggesting a core soil microbiome. In the SL+OM, the Actinobacte-
ria were dominant. Members from this phylum are decomposers and have
the ability to utilize an array of nutrient sources which could be why they
were pervasive in this soil rich in organic matter (Wang 2016). Another

common bacterial phylum present was Acidobacteria, which are characteristically predisposed to tolerate acidic conditions. Temperature was an integral driver in the shifts observed for Acidobacteria. Overall, both machine learning and relative abundance metrics revealed that temperature changes would be the determining factor to predict the community response to disturbance. This information is imperative to accurately predict microbial community dynamics in response to environmental variables.

3.9 Conclusion

In both the classification analysis and class level taxa analysis, the soils surveyed demonstrated that the influence of temperature and moisture was critical, though not equivalent in its impact. The classification algorithms showed that temperature was the most accurate variable for classification. However, when examining the significant taxa at the class level, there appeared to be dynamic changes due to temperature and moisture. These observations suggest that there was an interplay between the two variables that provide conditions for particular microbial populations. The extent of these shifts should be further explored. Additionally, organic matter content played a role in the microbial response to particular conditions. The soils without organic matter (L-OM and SL-OM) were noticeably impacted by temperature and moisture specifically in the emergence of underrepresented taxa which could suggest that lower organic matter promoted more specialized strategies in the soil communities. Temperature and moisture are indeed driving forces behind community composition and can therein be traced back to the varying temperature regimes. With more research, it may be possible to further classify the unique microbial communities that specialize at each temperature and moisture step and begin utilizing the community makeup to more accurately validate the environmental conditions they experienced.

4 CRREL Soil Microbiology Mathematical Model of Soil Activity

The objective of this task was to model the respiration and community composition datasets generated from the empirical incubation study described in Section 3. Understanding the behaviors generated by organisms in soils is a complex phenomenon (Paul 2014) requiring connectivity between different organisms, soil topology, and ambient weather attributes. Many studies have looked into the temperature dependence of soil respiration, often referred to as Q_{10} and its appropriate relationship to both temperature and respiration (Fang and Moncrieff 2001; Lloyd and Taylor 1994; Tjoelker et al. 2001). The non-dimensional sensitivity measure Q_{10} is a constant round value of 2 but increases exponentially with respect to temperature (Fang and Moncrieff 2001). Moreover, another variable that has been of interest is microbial community composition as measured by diversity of bacteria or fungi in the soil. In soil, Shannon and Simpson diversity indices have been used to measure diversity, which is the richness and evenness of organisms (Grice et al. 2009). We sought to determine relationships between Q_{10} and the diversity indexes. We explored the empirical dataset of soil activity in Barbato et al. (2015) by calculating the temperature sensitivity index Q_{10} and study how it is related to diversity indexes (Shannon and Simpson) calculated from phylum-level sequencing datasets described in Section 3 of this report. In addition, clear statistical approaches to do such a comparison and the limitations encountered in our study are provided.

4.1 Is the relationship between Shannon diversity H and Q_{10} linear?

In what follows, we used a calculated Q_{10} value with the formula:

$$Q_{10} = \left(\frac{R_2}{R_1}\right)^{\frac{10}{T_2-T_1}}$$

and Shannon diversity index H

$$H = -\sum_{i}^{n} p_i \, ln(p_i)$$

with p_i being the proportion of the bacteria/fungi i and n is the total number of bacteria/fungi. The objective of this research was to establish a relationship between Q_{10} and H.

First, we transformed the data and proceeded to calculate the Q_{10} using the linear regression formula:

$$Ln(Y) = \alpha T + \beta \qquad \text{and} \qquad Q_{10} = e^{10\alpha}$$

as illustrated in Pavelka et al. (2006). Note that "Ln" is the natural logarithm and "T" is for temperature. The SL+OM soil was used as an example and the results are presented in Table 5.

Table 5. Values of R², α, Q10, H with respect to change in temperature for soil SL+OM.

	Very Low Moisture	Low Moisture	Medium Moisture	High Moisture	Very High Moisture
R²	0.91699	0.9799	0.9604	0.9401	0.9475
α	0.0599	0.0635	0.0567	0.07918	0.06437
Q10	1.82	1.88	1.76	2.21	1.9
H	1.964	1.766	1.750	1.811	1.820

The Q_{10} calculated corresponded to high values of R^2 which illustrates the accuracy of our result based on the formula above (Table 5). Each moisture level had a slightly different α, which is the coefficient in the relationship between temperature and respiration. In addition, all Q_{10} values ranged from 1.7 to 2.2, as reported in the literature. Interestingly, under the High Moisture condition, a 10 °C change in temperature resulted in the greatest change in respiration/soil activity.

The diversity index (H) values were calculated for both bacteria and fungi grouped by moisture condition as well. Therefore, five H values were obtained within each soil group. Again, the SL+OM soil served as an example and yielded bacterial diversity values ranging from 1.75 to 1.964 (Table 5). Interestingly, the low moisture soil had the highest diversity.

Now that Q10 and H values were acquired for each soil group, a Spearman correlation was used to test the relationship between Q_{10} and bacterial diversity (H). For SL+OM, the correlation was 0.3 (p= 0.68). For SL, the

correlation was 0.4 (p= 0.52). For L-OM, the correlation was 0.1 (p= 0.95), and for SL-OM, the correlation was -0.5 (p= 0.45). Spearman between Q10 and fungal diversity was also calculated. For SL+OM, the correlation was 0.6 (p= 0.35) and for SL, the correlation was 0.3 (p= 0.68). For L-OM, the correlation was 0.4 (p= 0.52) and for SL-OM, the correlation was -0.9 (p= 0.08).

While some of the correlation coefficients were relatively high (i.e. > 0.6), the Spearman correlation showed that there was not a linear relationship between Q_{10} and H (for bacteria and fungi) at a confidence interval of α=0.05. Therefore, we tested the possibility of a nonlinear relationship between Q_{10} and H through visualization and nonparametric fitting.

4.2 Nonlinear relationship between microbial diversity H and Q_{10}

Linearity was tested by plotting Q10 by diversity. All soils tested followed different nonlinear patterns for both the bacteria (Figure 15) and the fungi (Figure 16). Therefore the relationship between Q_{10} and bacterial diversity was nonlinear and such nonlinearity varied with respect to soil groups. This result was expected given that no clear pattern emerged from the bacterial and fungal diversity given the changes in temperature and moisture. Moreover, the observed nonlinearity did not take any known form; thus it could not be established parametrically.

Figure 15. Modified Q10 values (as presented in Table 5) with respect to the Shannon diversity index H of bacteria for all soil types. The red line indicate a spline interpolation of the data points.
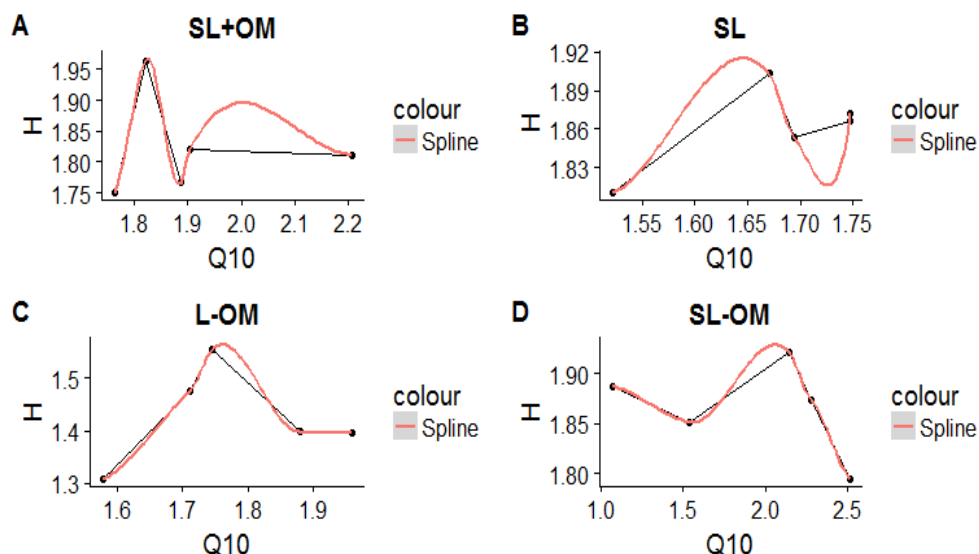
Figure 16. Modified Q_{10} values (as presented in Table 5) with respect to the fungal diversity index H for all soil types. The red line indicate a spline interpolation of the data points.
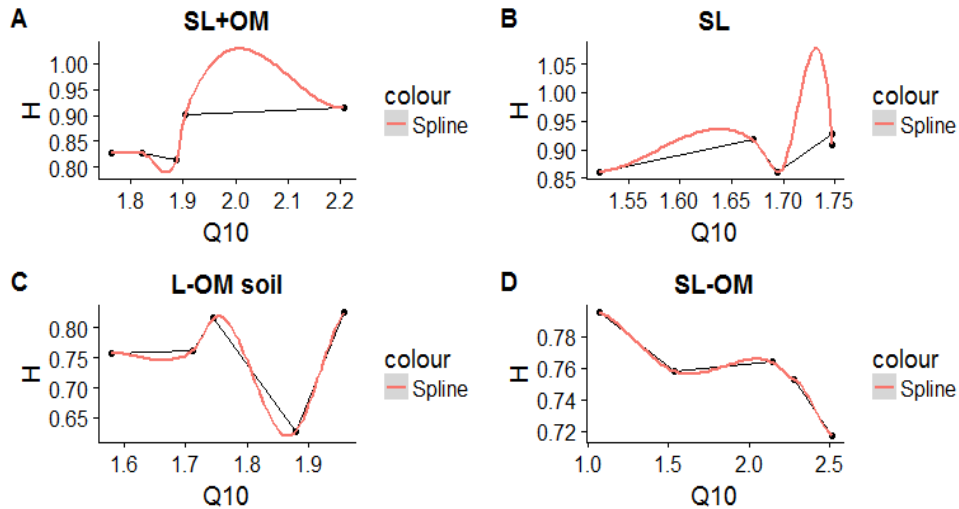
4.3   Interaction effects between bacterial and fungal diversity

In the work of Rousk et al. (2010), it was found that the relative abundance and diversity of bacteria were positively related to pH while the relative abundance of fungi were unaffected by pH. In order to determine how bacteria and fungi diversity are related, we investigated the relationship between the Shannon diversity of bacteria and fungi through data visualization with spline interpolation. In addition, through a spline interpolation, the interaction effect of the Shannon diversity of bacteria and fungi on Q10 was explored. A linear relationship did not exist between bacterial and fungal diversity (Figure 17). Furthermore, the nonlinear relationship observed varied with respect to soil groups. Bacterial and fungal diversity exhibited a positive correlation for the SL and SL-OM soils (Figure 17).

Figure 17. Plot showing the Shannon diversity index H of bacteria against fungi for all soil types. The red line indicate a spline interpolation of the data points.
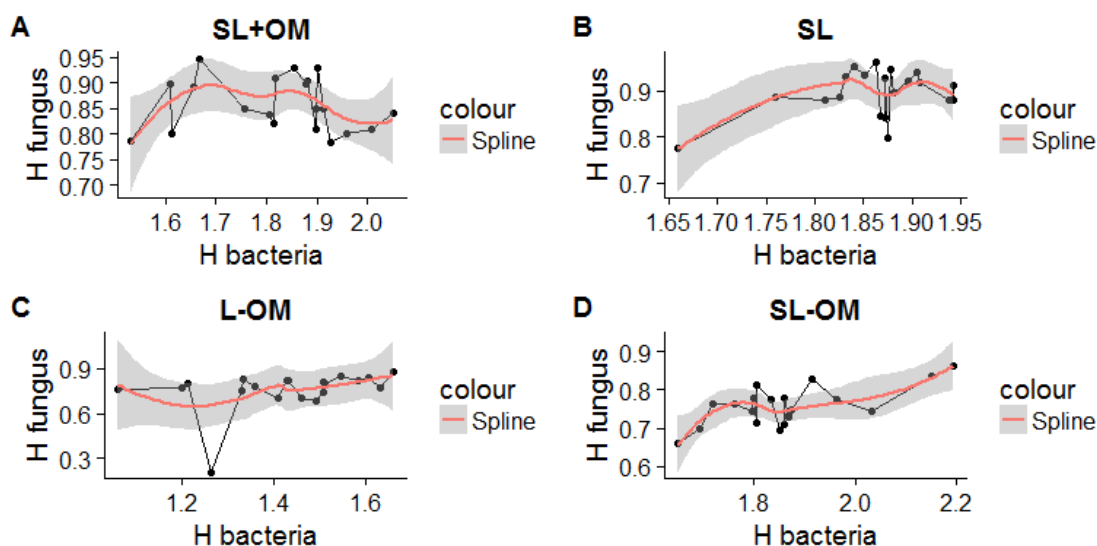
A linear relationship did not exist between the Shannon diversity H of the bacteria and fungi, as tested through Spearman correlation. For SL+OM and SL, the correlations were 0.7 (p= 0.23) and 0.3 (p= 0.68), respectively. Also, for L-OM, the correlation was 0.2 (p= 0.78) and for SL-OM, the correlation was -0.7 (p= 0.23).

Notice in Figure 18 that the relationship between Q_{10} and the interaction of Shannon diversity index for bacteria and fungi is not linear or monotonic. While the correlation was shown to be high, such values do not have any explicit meaning. This result showed the importance of verifying the monotonic or linear relationship assumption when calculating the correlation coefficient. No clear relationship could be obtained from the diversity (H) of fungi and bacteria or Q_{10} and the interaction of diversity indexes of bacteria and fungi. A further analysis must be performed to establish such result.

Figure 18. Plot showing the interaction effect of the Shannon diversity index H of bacteria and fungi on Q10 for all soil types. H_f*H_b on the y-axis denotes the interaction and the red line indicate a spline interpolation of the data points.
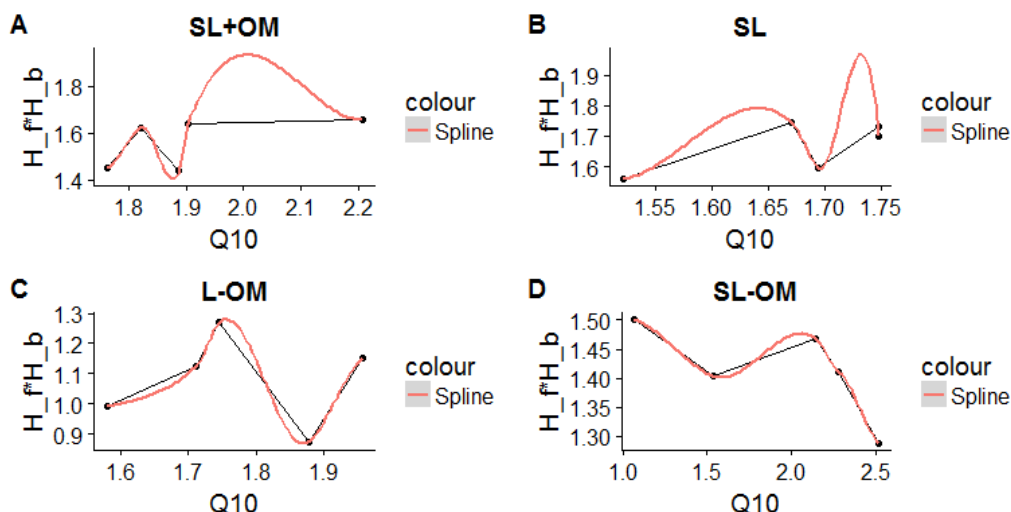
4.4 Relationship between Q_{10} vs. change in Shannon diversity index ΔH

Considering that Q_{10} and the diversity index have a one to one relationship, we attempted to transform the diversity indices to make them comparable. Note that Q_{10} is a rate and dimensionless while diversity index is the opposite. Therefore, we converted the diversity index into rates by taking the present value and dividing by the previously calculated value. Diversity was converted into a rate by performing $\Delta H = \frac{H_{j+1}}{H_j}$ where j represents the temperature index across the relative abundance of the microbes. For instance, from the diversity value obtained at temperature 5 °C and 15 °C, $\Delta H = \frac{H(15)}{H(5)}$ in the same way that $Q_{10} = \left(\frac{R_2}{R_1}\right)^{\left(\frac{10}{T_2-T_1}\right)}$. Now the diversity ΔH is dimensionless and has a one to one relationship with Q_{10}. A linear relationship did not exist between Q_{10} and the change in Shannon diversity index ΔH for both bacteria and fungi (Figure 19 and Figure 20). Based on these findings we continued the investigation using the Simpson diversity index (D), which takes the relative abundance of the organisms into account, giving more weight to more dominant ones.

Figure 19. Q10 vs. the change in Shannon bacterial diversity. The red line indicates the spline interpolation of the data.
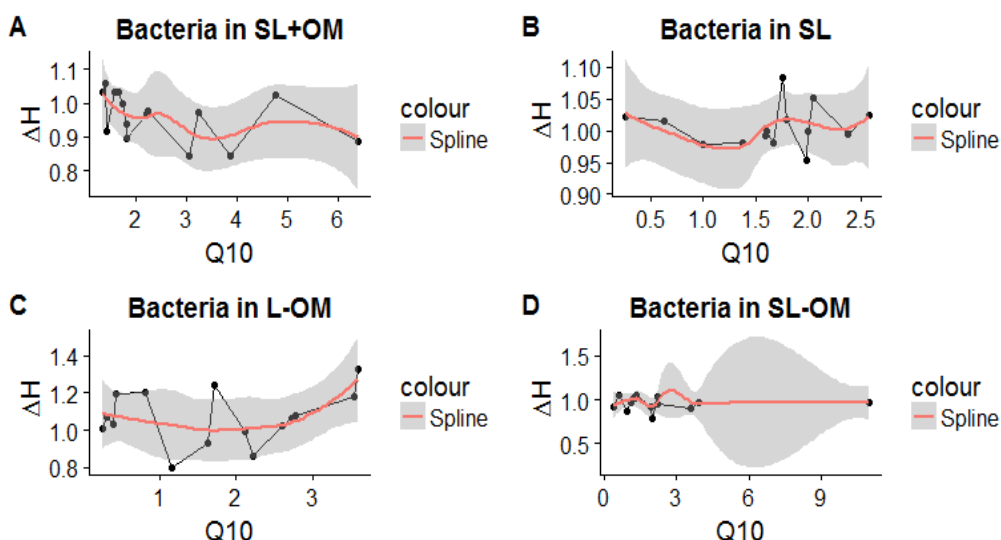
Figure 20. Q10 vs. the change in Shannon fungal diversity. The red line indicates the spline interpolation of the data.
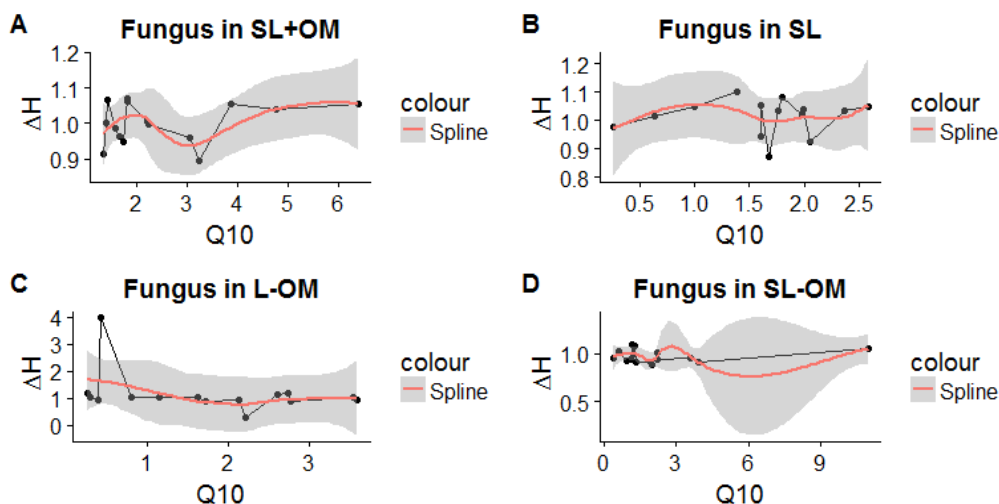
Simpson's diversity was calculated using the following formula.

$$D = \frac{1}{\sum_{i=1}^{S} p_i}$$

We then changed the diversity index into rates by taking the present value of D divided by the previous calculated value the formula is denoted by $\Delta D = \frac{D_2}{D_1}$. We obtained the change in D and compared it to Q_{10}. The results for the Simpson diversity followed a similar trend as the Shannon diversity, but with different magnitude. There they were not presented.

4.5   Relationship between the CO_2 efflux (respiration rate) and Shannon diversity index

After exploring the existing relationship between Q_{10} and the diversity index, the results indicated that there was not a linear relationship between these two variables. Also, the nonlinear relationship had not been established. In what follows, we investigated the relationship between the Shannon diversity index, H, and respiration rate.

The Spearman correlation was performed to study the monotonic relationship between the respiration rate (CO_2 efflux) and Shannon diversity index. For the soil SL+OM, the correlation was -0.66 (p= 0.002). For SL, the correlation was 0.37 (p= 0.1076). For L-OM and SL-OM, the correlations were 0.49 (p= 0.02) and -0.36 (p= 0.12), respectively.

Figure 21. Plot of the CO_2 efflux (respiration rate) vs. the Shannon diversity in bacteria. The red line indicates the spline interpolation of the data.
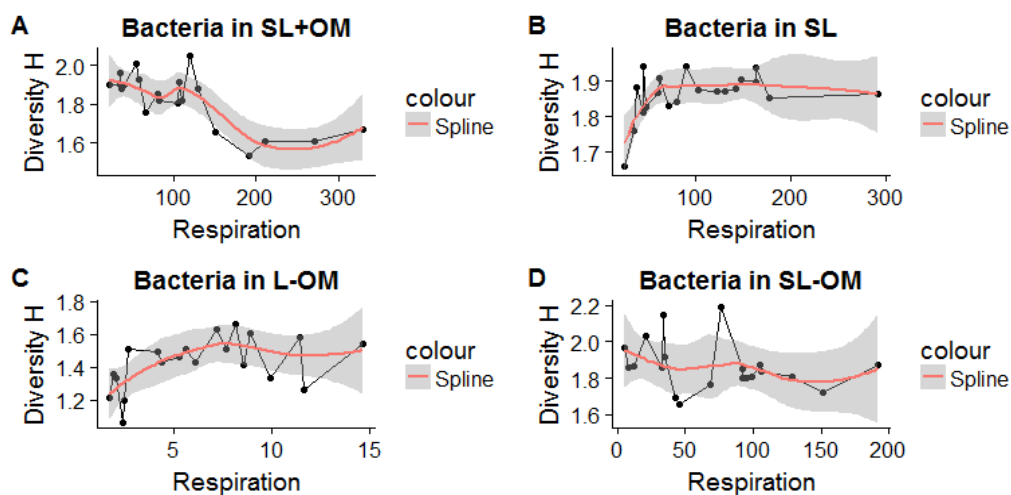
A nonlinear relationship was observed between diversity and respiration for all soils (Figure 21). Such relationship follows a logistic pattern for SL (Figure 21B) and L-OM (Figure 21C). The SL+OM exhibited a slight exponential decay pattern (Figure 21A). However, a rigorous statistical analysis must be done to establish such a functional relationship and this has not yet been explored in our analysis. In addition, Spearman correlation revealed a monotonic relationship between respiration and diversity H only in SL+OM and L-OM. Such result suggests the importance of considering soil properties when modeling microbial activity within different soils.

4.6 Discussion

Soil samples used in our study differed by their textures (e.g. silt, clay, sand percentage) and chemical properties (magnesium, calcium, potassium, etc.). These properties can alter the microbial response under different temperatures and moisture content. We used several different methodologies to test whether a relationship existed between the microbial diversity and the temperature sensitivity index Q10. Two diversity indexes were used in our analysis: Shannon diversity, denoted H, and Simpson diversity, denoted D. Both diversities were calculated for bacterial and fungal communities while Q_{10} remained the same as it pertains to the rate of change of a biological system with respect to an increase in temperature by 10 °C.

A correlation test was performed first. A monotonic dependence did not exist between Q_{10} and both diversity indexes H and D. In addition, a spline interpolation between Q_{10} and diversity index H indicated that soil properties affect the microbial response to temperature and moisture. Q_{10} was calculated using a linear regression formula as illustrated in Pavelka et al. (2006). A Spearman correlation between the diversity indexes and the newly calculated Q_{10} revealed that no linear relationship existed between the variables. This result again emphasizes the presence of soil properties affecting the relationship between Q_{10} and the diversity index. Furthermore, the relationship between respiration rate and microbial diversity may be inverse or proportional, depending on the characteristics of the soil. This was corroborated by the Spearman correlation.

4.7 Conclusion

Several methods were explored to reveal the relationship between Q_{10} and diversity indexes (mathematical measures of the diversity of a species in a community). From the analysis, the relationship between Q_{10} and diversity followed a nonlinear pattern. However, the nonlinearity was dependent upon the topology of the soil, ambient temperature, and other variables that may not have been taken into account in our analysis. While the omitting variable may be a limitation of our work, we believe this may be the first analysis in the direction of establishing a robust relationship between Q_{10} and species diversity. Determining the soil attributes that played strong roles in the relationship between microbial diversity and Q_{10} could lead us in the direction of discovering which species can be present under what conditions. This should be the focus in future studies.

5   CRREL Soil Microbiology DRTSPORE Platform and Validation

The aim of this task was to create a visualization platform to display soil activity predictions in a geospatial format. The conceptual flow of information for the platform is highlighted in Figure 22. External terrain data such as soil type and environmental predictions such as soil temperature and moisture serve as input parameters to the platform. Mathematical models described in Section 4 are then invoked to develop predictions of soil function, which is soil activity in this use case. The output soil function data is then displayed over a geospatial map in a red-amber-green color scheme for the area of interest selected by the user.

Figure 22. Conceptual flow of information for DRTSPORE platform.
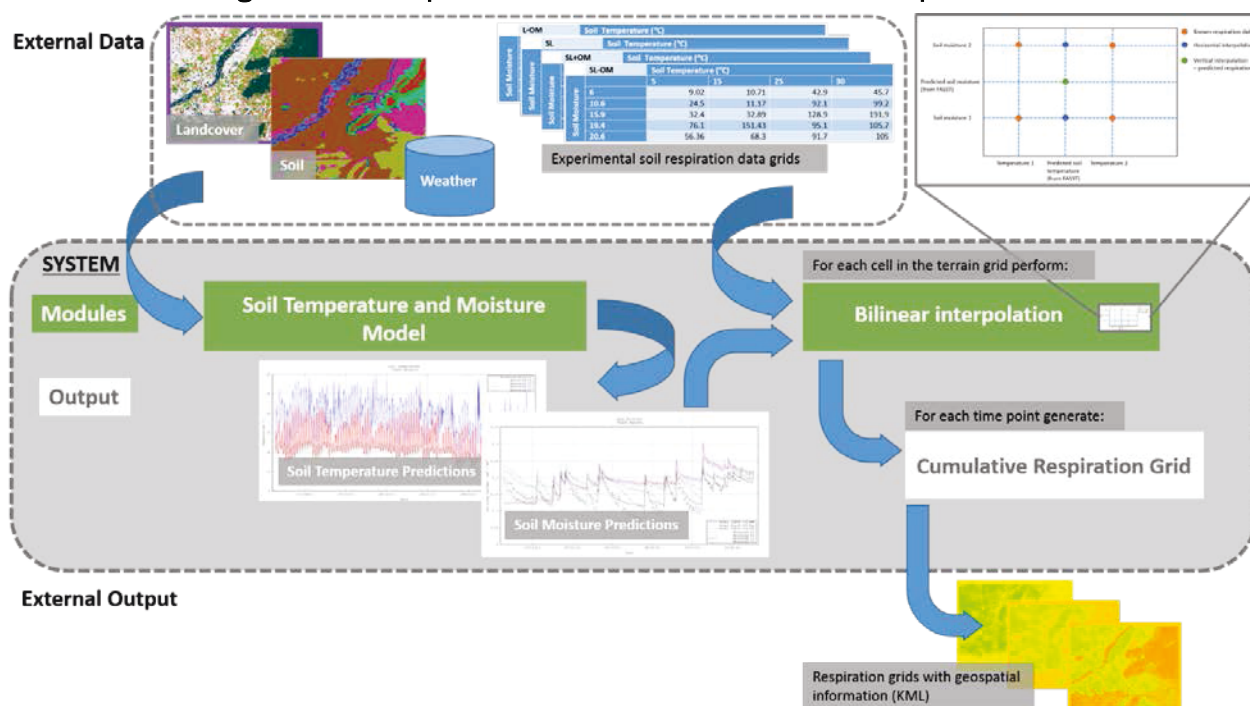
In the graphical user interface (GUI) in the ArcMap toolbar, the user selects the area of interest, date/time, output folder, mathematical model (through a look up table (LUT) folder), and the pixel size of the preferred grid space (Figure 23).

Figure 23. GUI of the DRTSPORE platform in ArcMap.

Once the AOI has been selected, the platform reaches to the Creare, Inc GeoWatch Tool Webservice and obtains information regarding soil texture, soil moisture, and soil temperature. These are current datasets or predictions that are forecasted up to three days in advance. There are 16 soil classifications available in the GeoWatch Tool aligning with the United States Department of Agriculture classifications. We have soil respiration models for four of these soil types and have grouped our four to the remaining 12 classifications available. This is an area where the DRTSPORE Platform can grow significantly in the future. By including more soil datasets, we will have a better fundamental dataset from which to draw more accurate predictions. The soil temperature and moisture data are then fed into the soil microbial respiration model which supplies a respiration rate. These rates are calculated per the grid space selected by the user and are then populated onto the map.

We used an example to demonstrate the platform that includes an AOI near CRREL (Hanover, NH, USA) that encompasses Vermont and New Hampshire. The soil type map for this AOI is displayed in Figure 24A. In this particular area, five soil textures were represented. In the GeoWatch Tool, sixteen soil textures were available. Then, the DRTSPORE platform

received soil temperature (Figure 24B) and soil moisture (Figure 24C) predictions based on the time-stamp selected in the GUI (Figure 24). The values were ingested in the DRTSPORE linear mathematical model and the final output was soil respiration rates displayed on the map (Figure 24D). The soil respiration rates were translated to a color symbol of red-amber-green whereby the green values reflected higher soil activity (soil respiration rates) and the red values reflected lower soil activity (Figure 24D). A legend denoting the activity level is critical to place the information into a frame of reference when comparing multiple areas or the same area over multiple time stamps.

Figure 24. Output of DRTSPORE platform in ArcMap. A) Soil texture, B) soil temperature, C) soil moisture, and D) predicted soil activity of the AOI.
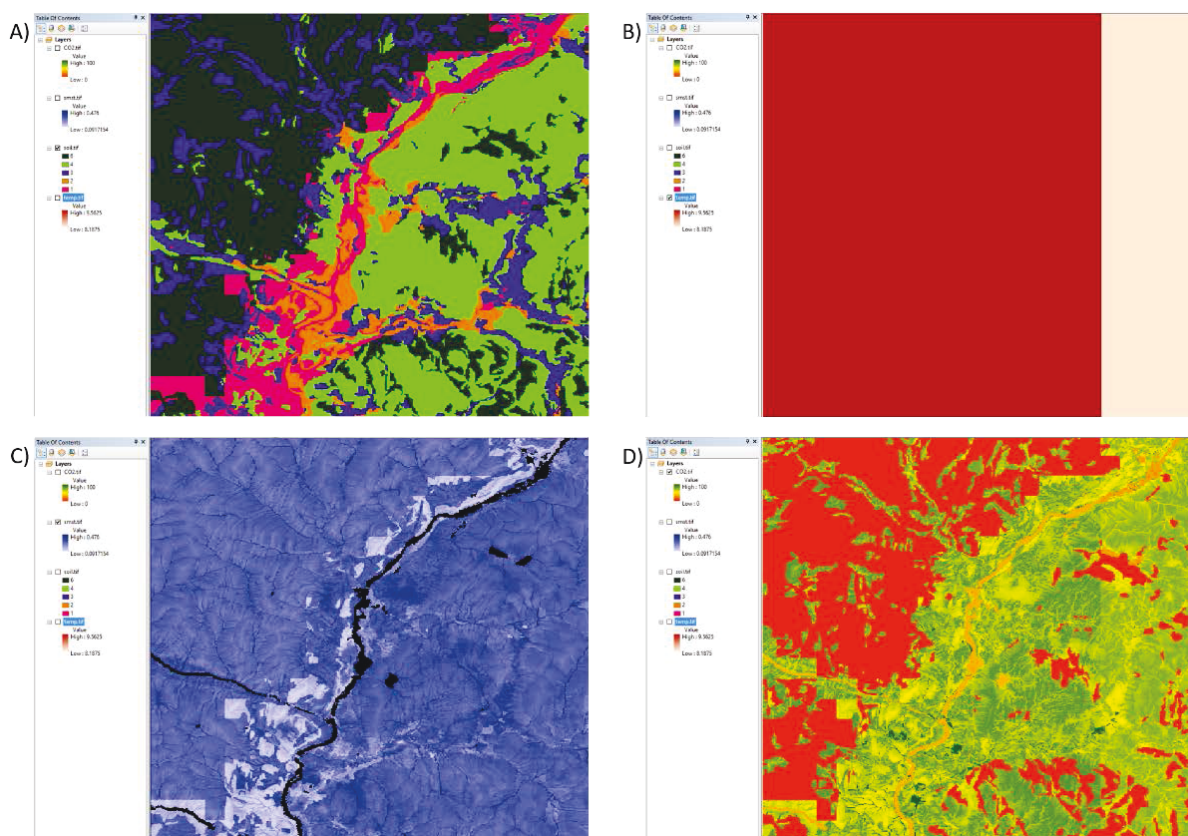
6 Validation of Soil Activity

6.1 Method

Field microbial respiration measurements were started at the CRREL field site on 25 September 2017 using an ADC BioScientific LCi-SD field respirometer. Each lane contains four measurement areas separated from the surrounding soil by polyvinyl chloride (PVC) collars. Measurements were collected from each replicate of all lanes at least once per week until snow began to fall. Measurements occurred at no set time and on different days of the week, generally during periods with the clearest weather for that week.

The field respirometer was operated using an area setting of 182.49 cm² which is equivalent to the area inside the PVC collars. Airflow rate was set to 200 $\mu mol \cdot s^{-1}$. For each measurement, the field respirometer leaf chamber was securely sealed to the PVC collars using a custom made aluminum ring with a rubber seal. Once sealed, the operator waited at least 20 minutes on each repetition to allow microbial generated CO_2 to stabilize inside the chamber before recording CO_2 and temperature readings. Temperature was measured at a depth of ~2.5 cm below the soil surface near each repetition using the temperature probe accessory attached to the field respirometer. Additional temperature and volumetric water content (VWC) measurements were recorded by in situ probes installed at the CRREL field site, at depths of 2.5 cm, 5 cm, and 10 cm below the soil surface, described in Section 2.

Starting on 3 Oct 2017, measurements were collected with the LCi-SD's automated logging feature enabled. The automated logging feature records all measurements the field respirometer takes, including ΔC where $\Delta C = (CO_{2atmospheric} - CO_{2leaf\,chamber})$. Both values are measured in units of μmol (of CO_2) $\cdot mol^{-1}$(of atmosphere) as well as soil and air temperature. Measurements were recorded every 48 seconds over the 20 minute measurement interval for each repetition, then saved to the local memory card on the LCi-SD in a CSV format.

Over the course of the field measurements, moss grew inside the PVC collars on all four lanes, and likely affected the CO_2 concentration readings in the leaf chamber to an unknown degree. Moss was removed on 9 Nov 2017 and measurements resumed one week later. At the time of removal, moss

in the SL and L-OM lanes covered most (>90%) of the soil surface inside the PVC collars. Several measurements with extremely positive ΔC values taken during the first week were later considered to be the result of an error with the LCi-SD. When the fan inside the leaf chamber failed, the LCi-SD began to record extremely positive ΔC values accompanied by an error message. Although no error message was noted for these values, it is assumed that this was due to user error. Graphs for which these values have been removed based on the criteria above were marked with "adjusted".

6.2   Results

Over the course of the fall 2017 season, ΔC values generally became more positive over time. This suggests that microbial respiration may largely be controlled by temperature variability, as the soil temperature generally decreased towards the winter (Figure 25). However, this downward trend may also be related to the increase of moss over time and a resulting increase in photosynthetic respiration. As noted above, the L-OM lane in particular seems to be an outlier, with ΔC values suspected to be responding more to weather conditions/time of day and the associated changes to photosynthetic respiration due to moss.

The ΔC values measured by the LCi-SD were also compared with soil temperatures from both the respirometer and in situ probes, as well as VWC of the soil as measured by the permanent probes. These probes recorded a temperature value every 15 minutes. The values were taken at multiple depths along the site and averaged. Because the temperature probe on the respirometer recorded close to the PVC collar and at the same time as the ΔC values, it likely provided a more accurate temperature localized to each individual repetition. For this reason, the following ΔC value vs. temperature graphs include only temperature as measured by the respirometer.

The ΔC values in all lanes were less correlated with temperature after the moss was removed. At this point, there was only one week of measurements taken before the soil began to reach freezing temperatures. On sunny days during this colder period, the top ~1 cm of soil tended to thaw, creating a slick, muddy top layer, with generally more negative ΔC values, indicating increased microbial respiration. However, temperatures measured by both the respirometer and in situ probes were taken below this layer and were typically close to or below freezing. Therefore, there was a disconnect between ΔC values and soil temperature because the respirometer measured surface soil respiration where there was daily soil thaw and

the temperature probe was deeper where the soil may have remained close to freezing. For example, the automated log at the SL+OM lane for one sunny day – on 20 Nov 2017 – shows a constant soil temperature despite a cessation of microbial activity. A closer look at air temperature recorded by the field respirometer in the same log shows that the air temperatures, and therefore likely soil surface temperatures, dropped over the same period. In this example, it seems clear that the respirometer was measuring soil temperatures below the layer of soil where microbes were most active. Therefore, it was not representative of the soil temperature where most microbial activity occurred (which was likely warmer). The soil temperature of the top layer of soil during these periods, if it had been measured, may have correlated better with ΔC values.

Figure 25. Average ΔC ($\mu mol_{CO_2}/mol_{atmosphere}$) variability as measured at the soil lanes through time. More negative ΔC values indicate higher microbial respiration. Moss was removed on 09 November 2017.
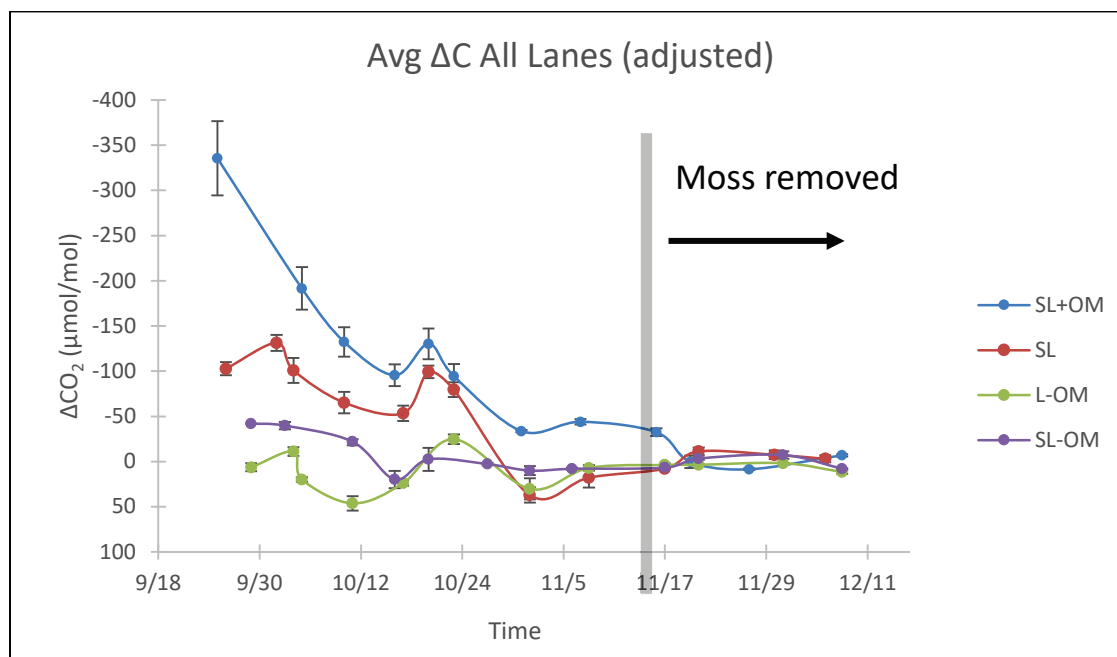
Figure 26. ΔC ($\mu mol_{CO_2}/ mol_{atmosphere}$) variability vs. soil temperature (°C at 2.5 cm below soil surface) of the soil lanes. More negative ΔC values indicate higher microbial respiration.
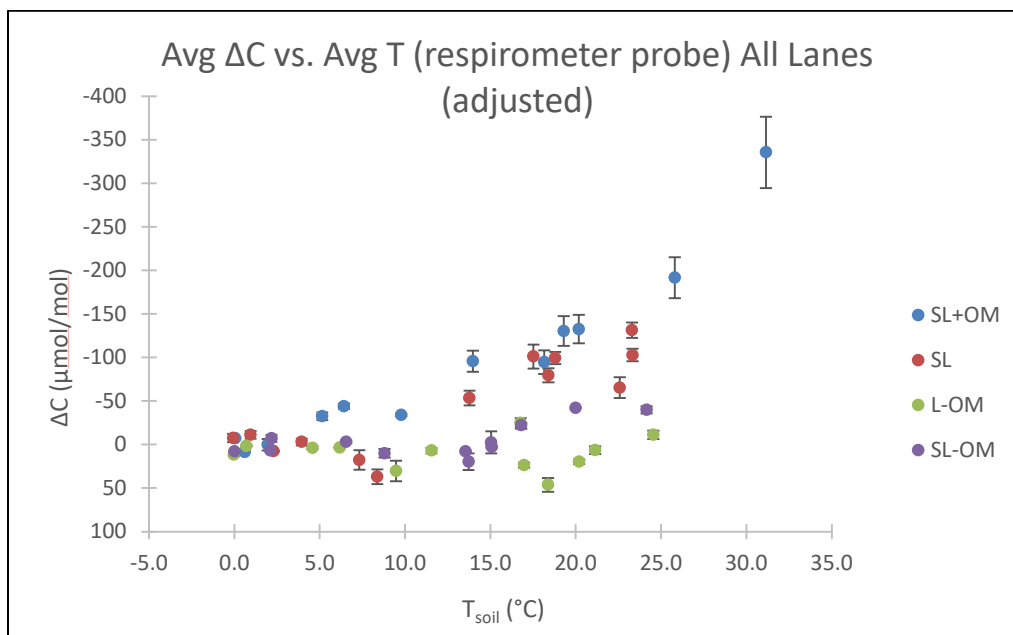
When ΔC values were compared to VWC, the expected bell-shaped curve from Moyano et al. (2013) was not observed (Figure 27). The ΔC values compared to VWC at the SL+OM lane appeared to come the closest of the four soil types, but still failed to yield the expected curve. It seems likely that the effect of photosynthetic respiration on ΔC was the biggest contributor to this disparity. In the future, maintenance to reduce photosynthetic activity is paramount. Higher VWC values also appeared to have generally occurred towards the end of the fall, around the time the moss was removed, which may have compounded the effect. The median VWC values for each lane also appeared to be inconsistent, this difference is likely due to the fact that soils with finer texture and/or organic matter tend to have a higher water capacity (Weil and Brady, 2016).

Figure 27. Average ΔC ($\mu mol_{CO_2}/ mol_{atmosphere}$) vs. average VWC (m³/m³ at 2.5 cm below soil surface) of all lanes. Error bars indicate the standard error of averaged ΔC values. More negative ΔC values indicate higher microbial respiration.
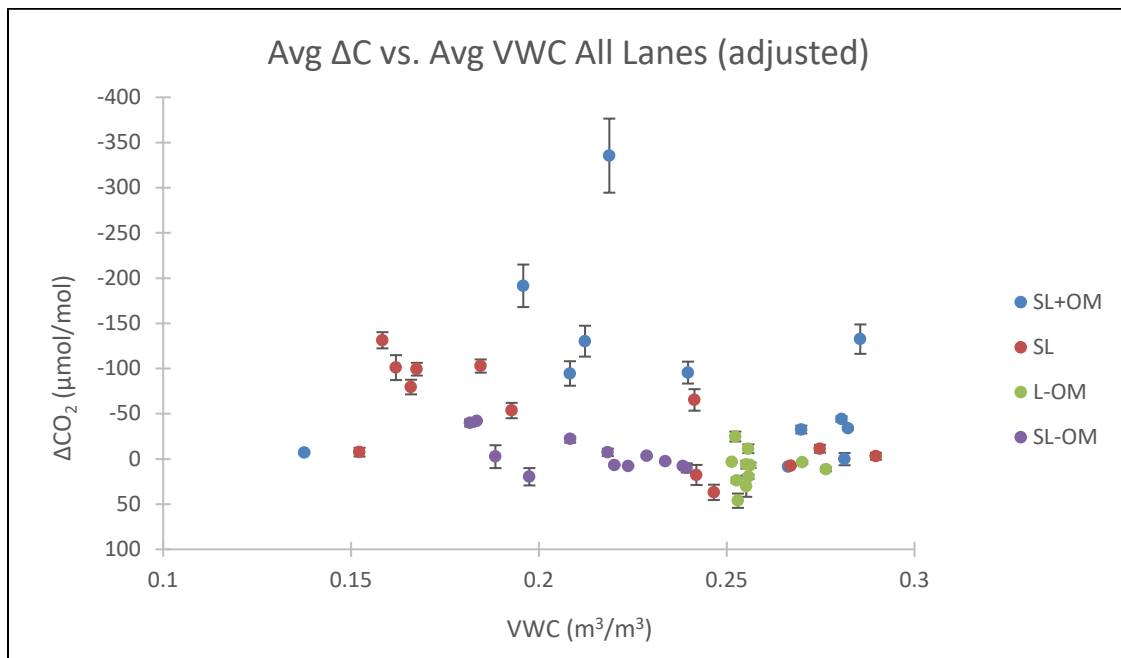
7   CRREL DRTSPORE Application: Microbial Fuel Cells

Available power is a consistent limitation to the Warfighter in theater. Therefore, there are ongoing efforts to develop smaller, longer-lasting batteries. The rapidly growing technology of microbial fuel cells (MFCs) is one example of a paradigm shift. MFCs are bioelectrochemical systems that convert energy contained in organic matter into electrical energy. They not only can provide electrical generation but also waste clean-up, via the microbial degradation of organic matter in the waste. From its inception over 100 years ago, the MFC technology development was very slow, and only recently accelerated, especially over the last two decades. The technology is now at an exciting point of showing great potential for clear cut practical and commercial applications. Our prior laboratory study showed that soil based MFCs could maintain performance at high and low temperatures (Barbato et al. 2017), considering their potential to be implemented under dynamic field conditions. However, there is still much effort required to take the technology from the laboratory to field scale.

The aims of this task were to: (1) understand physical and biological drivers in laboratory-based MFC performance and (2) adapt the laboratory MFCs for implementation at the CRREL Soil Micro Field Site described in Section 2. Research questions included which soil type produced the most power, whether Envirotac (a common soil strengthening polymer used by the Army) improved soil MFC performance, and could the MFCs be sustained in the field under highly variable environmental conditions (e.g., temperature and precipitation)?

7.1   Test polymer on MFC performance

Envirotac is a soil-binding polymer that has been used on roads for military purposes. We used it to line a section of the soil to limit water loss (Figure 28). In order to test this, each soil type was used to construct two sets of MFCs with and without polymer lining. Before the MFC was built, the polymer was set to dry and harden overnight. Figure 28 shows a photo of a MFC with polymer.

Figure 28. Photo of an MFC with polymer lining in a plastic bin. Sieved soil was used for the MFC and the soil surrounding was not sieved.
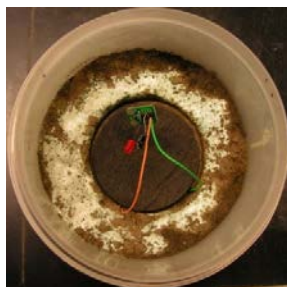
Polymer-lining increased power output of the MFCs, except for the L-OM MFC. The SL MFC power increased as high as 27 mW (Figure 29). By lining the soil with polymer, it is possible that the water exiting the MFC was minimized and the soil under the anode remained anoxic, which is key to optimal MFC performance. A microbial anode may involve strictly anaerobic bacteria, which are killed by the presence of oxygen. On the other hand, some electroactive bacteria can tolerate oxygen. In this case, the presence of oxygen diverts them from electron transfer to the anode and electrons are preferentially released to oxygen (Harnisch and Schröder et al. 2009). They prefer oxygen, their natural electron acceptor, over the electrode material. In supplement, oxygen can also support the development of non-electroactive bacteria (Qu et al. 2012), which may hinder electron transfer, resulting in low power output. Therefore, the polymer improved performance (Figure 29) and therefore was selected for the field implementation of the MFCs.

Figure 29. Power output of soil MFCs with and without polymer-lining.
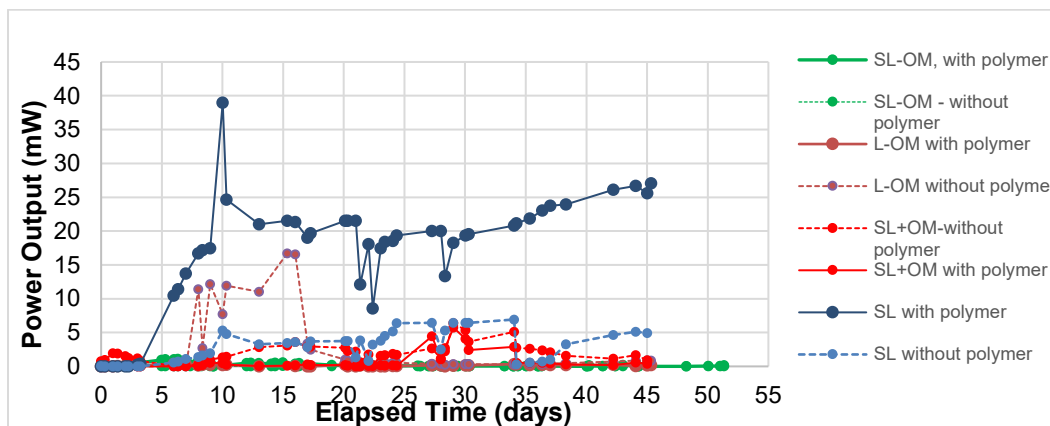
7.2   Adapt MFCs for the field

Soil MFCs were implemented at the CRREL Soil Microbiology Testbed to determine their performance under dynamic field conditions. Replicate soil MFCs were implemented in each soil lane in July of 2017 and measurements have been collected at different temporal frequency, from twice a day to a few times a month. Diurnal voltage changes were observed during the summer of 2017 (Figure 30). Morning readings were consistently lower than afternoon readings with an average voltage difference of up to + 17.7 mv. This strongly suggests environmental factors (e.g. temperature (air and soil), humidity, and wind) have a direct effect on MFC performance.

Figure 30. Summer diurnal voltage performance across all soil types for the period of one week.
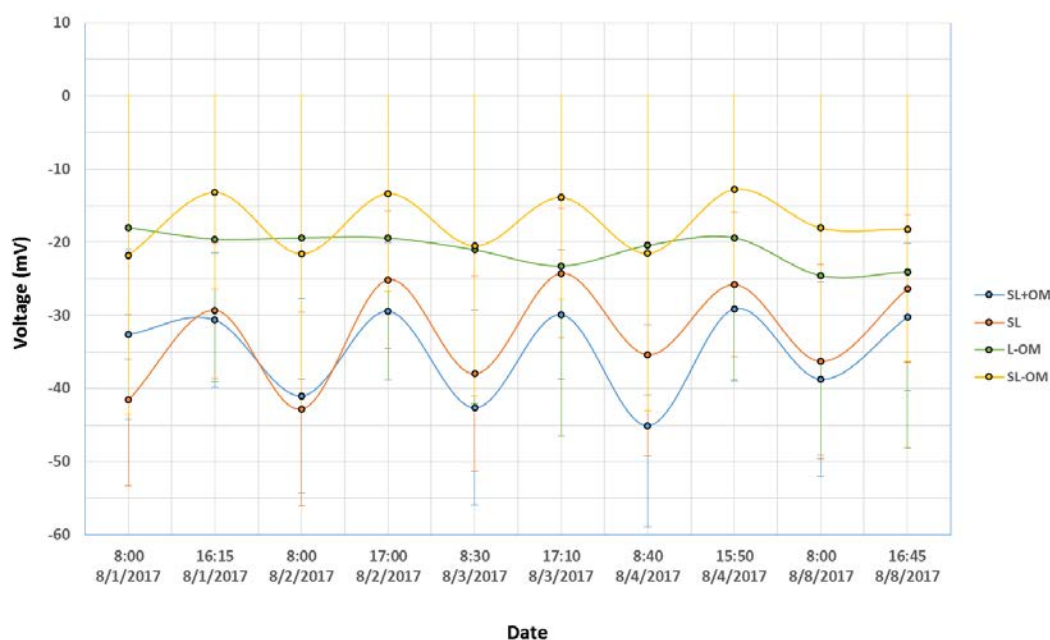
Voltage measurements have fluctuated over time, most likely due to weather and soil conditions, but the overall trend has been negative (Figure 31). During the initial week, the L-OM lane showed promising results averaging +20 mV with the highest positive voltage reaching +24.69mV on day twelve, while all other soil types trended toward more negative values. After 30 days in the field, all lane voltages grew increasingly negative. A visual inspection on day 60 revealed noticeably drier soils in all MFCs and most of the cathode felts were beginning to curl. On 7 September 2017, the surface cathodes of all MFCs were cleaned by removing

any loose debris from lane runoff and burrowing insects. To fix the cathode issue, toothpicks were used to pin them down to help maximize cathode-soil surface contact. Another important consideration was that the MFCs were losing water, resulting in aerobic processes at the anode. Therefore, 20ml of deionized water was added to the center of the anode soil, along with 60ml added to the center of the cathode soil to reestablish the moisture contents within these soils to their original values. After one week, all lane voltages began to stabilize and show a positive trend with the SL-OM soil performing the best (Figure 31). This is most likely due to the increased soil moisture and the renewed breathability of the cathode. Winterizing was performed on 11 December 2017 to prepare for the upcoming snowfall. This involved placing a small plastic liner over the top of the MFCs along with a weighted tin to prevent MFCs from being smothered with snow. Most soils generated stable voltages after winterization and through the winter months (December – March) except for the SL+OM soil, which grew increasingly negative throughout the winter. Spring thaw could have solicited more moisture infiltrating the MFC soil, and therefore, inducing a positive trend in voltage. Interestingly, it was also the SL+OM that showed a markedly positive increase during the spring thaw starting on 1 March 2018 (Figure 31). However, the SL, L-OM, and SL-OM soils sustained similar voltage values. In early to mid-spring, voltages began to show an increasingly positive trend in all soil types and then flattened during late spring (Figure 31).

Figure 31. Average voltage (mV) readings and trends for all field lanes from July, 2017 – June, 2018.
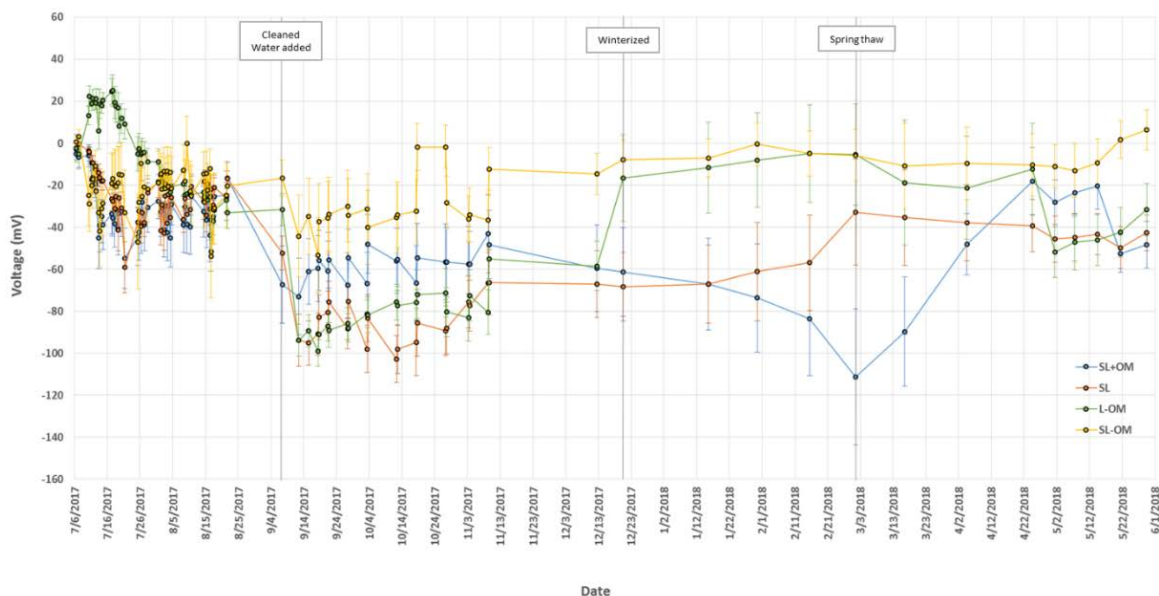
7.3   Future directions for field MFCs

Our previous experiments using soil MFCs incubated under controlled conditions in the laboratory have shown that a single MFC could reach a steady state in three days and produce enough energy to power a small LED light (Barbato et al. 2017). This shows promise for the future of terrestrial-based MFCs, and warrants more extensive testing in a field environment. Multiple MFCs could be wired together to produce the desired power. Most likely several MFCs will need to be clustered together to achieve optimal and sustainable power.

The negative trend in polarity could be caused by the loss of moisture content in the soil. As the battery remained in the ground, it was subjected to evaporation and loss of water through small cracks in the Envirotac potentially caused by the ground shifting or burrowing insects puncturing the lining. The addition of a more secure cover may provide a more controlled environment for the electrochemical reaction, keeping oxygen away from the anode. Despite these issues, the MFCs tested exhibited a strong negative voltage potential that could be solved by switching the polarity on the intended device or incorporating smart switching technology into the electronics themselves.

Further lab tests will be conducted using soils that have the greatest potential for microbial activity. So far, SL-OM and L-OM soils have been the yielded the greatest power output and therefore will be selected for in depth study. One key component of these experiments will be to isolate the factors that create a positive voltage potential. These factors may include soil moisture content, soil composition (organic matter, etc.), and alterations to MFC construction. Other experiments will include building MFCs on a larger scale. This would involve increasing the size and surface area of the cathode and anode and increasing the volume of soil for each. Theoretically, increasing the number of microbes on the surfaces of the cathode and anode would yield higher voltages.

8   Conclusion

Current geospatial models are limited because they do not incorporate biological processes that affect the behavior of soils. Organisms play a significant role in these soil behaviors and in turn control the attenuation of materials on surface soils, the recalcitrance of deposited asymmetric agents, and the mechanical properties and stability of soil. Therefore, their incorporation in modeling is paramount to more effectively predict Army relevant processes at the air-soil interface.

The DRTSPORE platform is a comprehensive decision aid that incorporates sophisticated weather models and terrain information to predict soil activity in a geospatial format. Different phases of the tasks are in various stages with the ultimate goal of improving the resolution and fidelity of the platform. It hosted six major tasks: CRREL Soil Microbiology Testbed for Soil Science Research, Empirical Data Collection, Mathematical Model of Soil Activity, DRTSPORE Platform, Validation of Soil Activity, and Sensor Technology Development.

For future development, we propose to significantly advance the DRTSPORE platform by adding high resolution data to predict specific soil functions and to incorporate Arctic and alpine soils to increase its representation of global soils. We will use our high resolution datasets developed from the Empirical Data Collection and DNA Tracking Technology tasks to quantify specific soil microbial functions such as the generation of exopolysaccharides for soil stability, the generation of electrons for advanced biosensors, and the degradation of deposited hazards for threat mitigation. These high resolution datasets will be used to mathematically model degradation rates and, by coupling knowledge of specific microbial generation and degradation pathways, we will increase our resolution to address the fate of military relevant compounds of concern. The models will serve as libraries that will be added to the existing graphic user interface. The improved environmental intelligence platform will provide Commanders a tactical decision aid to make better informed decisions about mobility, the placement and construction of a forward operating base which may be built by expeditionary forces, the placement of sensors, and the avoidance of areas where there is a potential for exposure to mobilized toxic materials in the soil.

References

Aronesty, E. 2011. ea-utils: Command-line tools for processing biological sequencing data. Durham, NC: Expression Analysis.

Barbato, R.A., K.L. Foley, and C.M. Reynolds. 2015. Soil temperature and moisture effects on soil respiration and microbial community abundance. ERDC/CRREL TR-15-6. Hanover, NH: U.S. Army Engineer Research and Development Center.

Barbato, R.A., K.L. Foley, J.A. Toro-Zapata, R.M. Jones, and C.M. Reynolds. 2017. The power of soil microbes: sustained power production in terrestrial microbial fuel cells under various temperature regimes. *Applied Soil Ecology* 109: 14-22.

Breiman, L. 2001. Random forests. *Machine Learning* 45(1), 5-32.

Caporaso, J. G., C. L. Lauber, W.A. Walters, D. Berg-Lyons, J. Huntley, N. Fierer, and N. Gormley. 2012. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *The ISME Journal* 6(8): 1621.

DeSantis, T.Z., P. Hugenholtz, N. Larsen, M. Rojas, E.L. Brodie, K. Keller, T. Huber, D. Dalevi, P. Hu, and G.L. Andersen. 2006. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Applied and Environmental Microbiology* 72: 5069-5072.

Fang, C. and J. B. Moncrieff. 2001. The dependence of soil CO_2 efflux on temperature. *Soil Biology and Biochemistry* 33: 155-165.

Grice, E.A., H.H. Kong, S. Conlan, C.B. Deming, J. Davis, A.C. Young, G.G. Bouffard, R.W. Blakesley, P.R. Murray, E.D. Green, and M.L. Turner, M.L. 2009. Topographical and temporal diversity of the human skin microbiome. *Science* 324 (5931): 1190-1192.

Harnisch, F. and U. Schröder. 2009. Selectivity versus mobility: Separation of anode and cathode in microbial bioelectrochemical systems. *ChemSusChem: Chemistry & Sustainability Energy & Materials 2*(10): 921-926.

Lloyd, J., and J. A. Taylor. 1994. On the temperature dependence of soil respiration. *Functional Ecology* 1:315-323.

Lozupone C, Knight R. 2005. UniFrac: a new phylogenetic method for comparing microbial communities. *Appl. Environ. Microbiol.* 71(12): 8228-8235.

Moyano, F.E., S. Manzoni, and C. Chenu. 2013. Responses of soil heterotrophic respiration to moisture availability: An exploration of processes and models. *Soil Biology and Biochemistry* 59: 72-85.

Paul, E.A. 2014. Soil microbiology, ecology and biochemistry. Academic press.

Pavelka, M., M. Acosta, M.V. Marek, W. Kutsch, and D. Janous. 2007. Dependence of the Q_{10} values on the depth of the soil temperature measuring point. *Plant and Soil* 29 (1-2): 171-179.

Qu, Y., Y. Feng, X. Wang, and B.E. Logan. 2012. Use of a coculture to enable current production by *Geobacter sulfurreducens*. *Applied and Environmental Microbiology* 78(9): 3484-3487.

Rousk, J., E. Bååth, P.C. Brookes, C.L. Lauber, C. Lozupone, J.G. Caporaso, R. Knight, and N. Fierer. 2010. Soil bacterial and fungal communities across a pH gradient in an arable soil. *The ISME Journal* 4(10):1340.

Tjoelker, M.G., J. Oleksyn, J. and P.B. Reich. 2001. Modelling respiration of vegetation: evidence for a general temperature-dependent Q10. *Global Change Biology* 7(2): 223-230.

Wang, C., D. Dong, H. Wang, K. Müller, Y. Qin, H. Wang, and W. Wu. 2016. Metagenomic analysis of microbial consortia enriched from compost: new insights into the role of Actinobacteria in lignocellulose decomposition. *Biotechnology for Biofuels* 9(1): 22.

Weil, R.R. and N.C. Brady. 2016. The Nature and Properties of Soils. Columbus EUA: Pearson Education.

REPORT DOCUMENTATION PAGE

1. REPORT DATE (DD-MM-YYYY)	2. REPORT TYPE	3. DATES COVERED (From - To)
September 2018	Final	

4. TITLE AND SUBTITLE

Dynamic Representation of Terrestrial Soil Predictions of Organisms'
Response to the Environment

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
ERDC 6.2 GRE ARTEMIS DRTSPORE

6. AUTHOR(S)

Robyn A. Barbato, Lauren Waldrop, Komi Messan, Robert Jones, Stacey J. Doherty,
Karen Foley, Christopher Felt, Michael Morgan, and Youl Han

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
U.S. Army Engineer Research and Development Center (ERDC)
Cold Regions Research and Engineering Laboratory (CRREL)
74 Lyme Road
Hanover, NH 03755

8. PERFORMING ORGANIZATION REPORT NUMBER
ERDC/CRREL TR-18-15

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Headquarters, U.S. Army Corps of Engineers
441 G Street NW
Washington DC 20314-1000

10. SPONSOR/MONITOR'S ACRONYM(S)
HQUSACE

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release. Distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT
The Dynamic Representation of Terrestrial Soil Predictions of Organisms' Response to the Environment (DRTSPORE) platform was created to *characterize* soil activities in barren, vegetated, and complex environments, *predict* biological-impacted processes in soil, and *generalize* the framework to solve a range of Army-relevant problems. Key biological processes include soil stabilization, emerging power sensors, brown out abatement, pathogen emergence, and contaminant and threat degradation. These capabilities are a new essential aid to national security, as they currently do not exist. The primary objective of this research effort, DRTSPORE, was to develop an environmental intelligence tool that adds a biochemical layer to current high resolution remotely sensed terrain and sophisticated weather model products. The mathematical models developed serve as libraries added to the existing graphic user interface. The improved environmental intelligence platform will provide Commanders a tactical decision aid to make better informed decisions about mobility, the placement and construction of a forward operating base, the placement of sensors, and the avoidance of areas where there is a potential for exposure to mobilized toxic materials in the soil.

15. SUBJECT TERMS

Soils – Remote sensing, Remote-sensing images, Soil biochemistry, Geospatial data, Geographic information systems

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**	UU	60	**19b. TELEPHONE NUMBER (include area code)**
Unclassified	Unclassified	Unclassified			