# A Functional Anova Approach to Detecting Changes in Soil Moisture and Temperature

Manju M. Johny

October 18, 2017

## 1 Introduction

Soil plays a pivotal role in maintaining rich and abundant biodiversity. Climate change poses significant challenges to the soil ecosystem, with profound implications for many aspects of life, from survival of microbial organisms to plant phenology [1] [4]. Yet, robust statistical approaches to precisely quantify the changes in various soil parameters has been lacking. Our study dissects the effect of climate change on abiotic aspects of soil ecology — soil moisture and temperature.

To mimic climate change, warmer temperatures and earlier snowmelt were experimentally simulated through the use of passive heating chambers and snow removal in the montane meadows[4]. Soil moisture and temperature measurements were recorded hourly for four treatments: (1) control, C; (2) heating, H; (3) heating + snow removal, HSR; and (4) snow removal, SR, with 3 replicate series in each treatment [4]. Our interest lies in establishing meaningful comparisons among groups of series. However, due to the presence of temporal dependence as well as the possibility of non-negligible correlations among series, the classical statistical inferential methods are not applicable. A classical one-way anova approach is inadequate here because it reduces the behavior of a full sequence of observations to a single mean value and does not account for the dependence structure. As a solution, a non-parametric permutation test was suggested to identify differences between groups of time series in our same experimental context by Sherwood et al [4]. While the permutation test considered the behavior of the entire curve, it is limited by the number of possible permutations given the small number of series in each group and fails to utilize the inherent functional nature of the series. The elements in each of our time series are realizations of a continuous process evaluated at discrete time intervals, and thus can reasonably be considered functional rather than as a vector of measurements. In this paper, we propose an anova test for functional data to detect the effect of climate change on soil ecology. The functional anova approach overcomes the limitations of classical one-way anova by considering features of the whole curve rather than just the mean, and provides a useful method of obtaining $p$-values through parametric bootstrap. Since many data can be considered functional in nature, the approach presented in this paper is widely applicable.

An overview of the contents of this paper are as follows. After giving a brief description of the data in Section 2, the smoothing procedure used for converting discrete data into functions is provided in Section 3.1. Section 3.2 gives an overview of functional anova, and Section 3.3 provides its practical implementation. The results are presented and discussed in Section 4, and a small simulation study regarding statistical power is discussed in Section 5. The major conclusions of our study are summarized in Section 6.

## 2 Data

The data consist of soil temperature and moisture measurements taken from a sagebrush meadow at an elevation of 2100 m in Grand Teton National Park, WY, and were collected by Dr. Diane Debinski's lab. Using a replicated block design, four treatments were applied: (1) control, C; (2) heating, H; (3) heating + snow removal, HSR; and (4) snow removal, SR. For each treatment, three replicates were established in 8x8 plots geographically: east, center, and west. Soil temperature was obtained at a depth of 5cm, and soil moisture was obtained at a depth of 25cm [4]. The measurements were taken hourly over a period of 5 months, from May 27, 2011 to September 27, 2011. In total, there were 2,954 time points for which temperature and moisture measurements were collected for each of the replicates in the four treatments. Some of these data are missing values, for which imputation is discussed at the end of Section 3.1. The raw soil moisture and temperature time series for each of the treatments are given in Figure 1.

# 3 Methods

## 3.1 Data Pre-processing

Although data may be functional by nature, it is typically collected as sets of discrete quantities. Thus, the first task in our analysis is to convert the discrete measurements of temperature and moisture into functional form. Since there is error associated with the measurements, the task of smoothing is vital for obtaining sets of purely functional curves that represent the behavior of the data without interpolation [3]. Smooth curves are constructed as a linear combination of $K$ independent basis functions, with careful consideration given to the choice of basis function. Let $x(t)$ be the functional representation of a series as a function of time, t. Then, the basis expansion of $x(t)$ is given by

$$x(t) = \sum_{k=1}^{K} c_k \phi_k(t) \tag{1}$$

where $\phi_k$ is the $k^{th}$ basis function. The choice of the bases functions depends on the underlying features of the data we wish to represent. Fourier bases are commonly used for cyclical data, and B-spline bases are commonly utilized for non-cyclical data. Examination of the data given in Figure 1 reveals cyclical pattern in temperature time series for which we choose a Fourier basis. Since the moisture time series lacks any such periodicity, we will use a B-spline basis. The number of basis, $K$, is often obtained by minimizing the Generalized Cross Validation (GCV) criterion developed by Craven and Wahba [6]. Since the replicate curves within a treatment can be considered realizations of a common process, we minimize the sum of the GCV obtained from the three curves within a treatment. For our data, complications arise from this minimization procedure because the GCV continually decreases as number of basis is increased. A basis number that is too large has high computational cost and overfits the data. We wish to obtain the optimal number of bases such that the general behavior of the data is retained in the functions without interpolating hourly variation. Thus, an acceptably smooth fit is achieved near the initial drop of the GCV criterion and is verified by visual comparison of the smoothed curve to the original data. As an example, Figure 2 shows the GCV minimization procedure for moisture time series in the heating treatment, and a resulting smooth curve from the treatment. Each of the moisture time series are represented in functional form as a linear combination of 120 B-spline basis of the third order. Each of the temperature time series are represented in functional form as a linear combination of 75 Fourier basis. Figure 3 gives the resulting functional curves for soil moisture and temperature post-smoothing.

The raw data contain approximately 0.7% missing values sporadically throughout the 12 moisture and 12 temperature time series, and all time points contained at least one replicate that had a non-missing value. While the missing values do not prevent the application of smoothing or the subsequent functional anova procedure, they pose problems with the GCV minimization procedure by producing NA values for the GCV criterion. In order to obtain the optimal basis number through the GCV criterion, we imputed the missing values in the following manner. For missing values in a time series with non-missing neighbors, we imputed by taking an average of its neighboring values. If a series contained a small sequence of missing values, we imputed using values from the closest replicate series if available. If the two replicate series were equidistant from the series containing missing values, we imputed by taking the average of the values from the two replicates. The GCV minimization procedure from the previous paragraph was applied to the data augmented with the imputed missing values, and was subsequently smoothed for analysis.

## 3.2 Functional ANOVA

An anova test for functional data (fanova) is used to test equality of mean curves between groups. In the classical one-way anova approach, equality of means is rejected when the between group variability is larger than the within group variability at a prescribed significance level. This idea is extended to the functional context in the form of an asymptotic test. We begin by giving an overview of the fanova test proposed by Cuevas et al.[2], and then present some modifications to the procedure.

Suppose we are interested in testing the equality of mean curves between $m$ independent groups. For $i = 1, \ldots, m$ and $j = 1, \ldots, n_i$, let $x_{ij}(t)$ represent the $j^{th}$ sample curve in the $i^{th}$ group as a function of time, $t \in [a, b]$. For each $i^{th}$ group, the $n_i$ sample curves may be considered realizations of a common $L_2$-process with mean $\mu_i(t)$ and covariance $K_i(s,t) = Cov(X_i(s), X_i(t))$. Additionally, define the functional sample mean curve of the $i^{th}$ group as

$$\bar{x}_{i.}(t) = \sum_{j=1}^{n_i} \frac{x_{ij}(t)}{n_i} \tag{2}$$

and the sample covariance of the $i^{th}$ group as

$$\hat{K}_i(s,t) = \sum_{j=1}^{n_i} \frac{[x_{ij}(s) - \bar{x}_{i.}(s)][x_{ij}(t) - \bar{x}_{i.}(t)]}{n_i - 1} \tag{3}$$

The null hypothesis we wish to test has the form

$$H_0 : \mu_1(t) = \mu_2(t) = \ldots = \mu_m(t) = \mu(t) \tag{4}$$

The following result, which may be used to test $H_0$, is presented by Cuevas et al. and a proof is provided in the same paper [2].

**Result: Asymptotic test for FANOVA (Cuevas et al. [2])**

Define $V_n$ as

$$V_n = \sum_{i<i'}^{m} n_i \|\bar{x}_{i.}(t) - \bar{x}_{i'.}(t)\|^2 \tag{5}$$

where $\|.\|$ represents the $L_2$ norm. Let $m$ be the number of groups, and $n = \sum_{i=1}^{m} n_i$ be the total number of sample curves. For $i = 1, \ldots, m$ and $j = 1, \ldots, n_i$, assume that

1. $n_i, n \to \infty$ such that $\frac{n_i}{n} \to p_i < \infty$.

2. $x_{ij}(t)$ corresponds to independent samples from a common $L_2$-process with mean 0 and covariance $K_i(s,t)$.

Then, under the null hypothesis, the asymptotic distribution of $V_n$ coincides with the distribution of $V$ such that

$$V = \sum_{i<i'}^{m} \|Z_i(t) - C_{ii'} Z_{i'}(t)\|^2 \tag{6}$$

where $C_{ii'} = \left(\frac{p_i}{p_{i'}}\right)^{1/2}$ and $Z_i(t)$ are independent Gaussian processes with mean 0 and covariance $K_i(s,t)$.

To test $H_0$, Cuevas et al. suggests the use of $V_n$ as the test statistic and obtaining the empirical distribution of $V$ through a parametric bootstrap procedure in which $Z_i(t)$ are generated from a Gaussian process with mean 0 and covariance $\hat{K}_i(s,t)$.

In this paper, we propose an alternative test statistic that preserves the distributional properties of $V_n$, and has the added advantage that it allows for direct visualization and uncertainty quantification of differences between the mean curves of each group. Our proposed test statistic $T_n$ is given by

$$T_n = \sum_{i<i'}^{m} \|\bar{x}_{i.}(t) - \bar{x}_{i'.}(t)\|^2 \tag{7}$$

A notable change in the form of the test statistic is the absence of the $n_i$ multiplier in its formulation. This modification has implications both for the simulation procedure and the asymptotic behavior of $T_n$. In particular, let $Z_i$ be the mean curve obtained from $n_i$ Gaussian processes having mean 0 and covariance $\hat{K}_i(s,t)$. Then, the reference distribution of $T_n$ is given by

$$T = \sum_{i<i'}^{m} \|Z_i(t) - Z_{i'}(t)\|^2 \tag{8}$$

In practice, using $T_n$ as the test statistic leads to analogous results as using $V_n$. However, the benefit of using $T_n$ is twofold. First, $T_n$ has an intuitive interpretation as the squared norm of the difference between pairwise sample mean curves. Second, the distances between the sample mean curves can directly be compared to the distances between the simulated $Z_i(t)$ resample curves. The latter, which was not possible in the Cuevas et al procedure, allows us an additional tool besides the $p$-value to visually understand the significance of differences between mean curves.

### 3.3 Parametric Bootstrap Procedure

We provide here the details of the implementation of the functional anova test through a parametric bootstrap procedure. For $i = 1, \ldots, m$ and $j = 1, \ldots, n_i$, recall that $x_{ij}(t)$ represents the $j^{th}$ sample curve in the $i^{th}$ group as a function of time, $t \in [a, b]$. For each $i^{th}$ group, the sample mean curve $\bar{x}_{i.}$ and sample covariance $\hat{K}_i(s, t)$ are given by Equation 2 and 3 respectively. For computation, the $x_{ij}$ and $\bar{x}_{i.}$ are rediscretized into a vector of length $T$. $\hat{K}_i(s, t)$ is obtained using the rediscretized sample curves as a $T$x$T$ sample covariance matrix. In the implementation of this procedure to our data, we used the number of basis functions involved in the smoothing procedure of our series as T. The following bootstrap procedure was then implemented for $N = 2000$ simulations.

1. Calculate the test statistic, $T_n = \sum_{i<i'}^{m} \|\bar{x}_{i.}(t) - \bar{x}_{i'.}(t)\|^2$, using euclidean distance to approximate the $L_2$-norm.

2. For each $i^{th}$ group, generate $n_i$ sample curves from a Gaussian distribution with mean 0 and covariance $\hat{K}_i(s, t)$. The mean of the generated sample curves in the $i^{th}$ group is denoted by $Z_i^* = (Z_i^*(t_1), \ldots, Z_i^*(t_T))$. The $Z_i^*$ are bootstrap resamples that approximate the continuous trajectories of $Z_i(t)$.

3. For $l = 1, \ldots, N$, simulate $Z_{il}^*$ as in the previous step.

4. For $l = 1, \ldots, N$, calculate $T_l = \sum_{i<i'}^{m} \|Z_{il}^* - Z_{i'l}^*\|^2$.

5. Calculate the $p$-value $= \frac{1}{N} \sum_{l=1}^{N} I(T_l > T_n)$.

The empirical distribution of $T_l$ approximates the asymptotic distribution of the test statistic, $T_n$. Hence, the parametric bootstrap procedure provides a computationally pleasing method for obtain $p$-values for functional anova tests. Additionally, the $Z_i^*$ represent the expected mean curves under the null hypothesis for each group. Thus, the N resamples of $Z_l^*$ can be plotted and compared to the sample mean curves from the data to visually understand the significance obtained from the $p$-value. This could be particularly useful if there are major shifts in the behavior of the process over time. Such changes may not be reflected in the magnitude of the $p$-value, but would include important scientific information.

# 4 Results and Discussion

The functional anova procedure described in Section 3.3 is applied separately to the temperature and moisture curves. Initially, we test for differences between the 4 treatment groups. Since there are only 3 sample curves in each of the 4 treatment groups, it is possible that this test may suffer from a lack of power. To mitigate this issue, two additional tests are performed in which we combine treatments such that the groups contain 6 sample curves instead of 3. The new groups are defined below.

First, curves are reassigned into two groups according to presence or absence of snow removal. The combined snow removal group ($\tilde{\text{S}}$) contains 6 sample curves from SR and HSR. The combined no snow removal ($\text{N}\tilde{\text{S}}$) group contains 6 sample curves from C and H. A functional anova test is then used to identify whether the combined snow removal group ($\tilde{\text{S}}$) is different from the combined no snow removal group ($\text{N}\tilde{\text{S}}$) for soil moisture and temperature.

Next, curves are reassigned into two groups according to presence or absence of heating. The combined heating group ($\tilde{\text{H}}$) contains 6 sample curves from H and HSR. The combined no heating ($\text{N}\tilde{\text{H}}$) group contains 6 sample curves from C and S. A functional anova test is then used to identify whether the combined heating group ($\tilde{\text{H}}$) is different from the combined no heating group ($\text{N}\tilde{\text{H}}$) for soil moisture and temperature.

The results of the functional anova tests on the moisture and temperature time series are presented in Section 4.1 and Section 4.2 respectively. In Section 4.3, we present our justification for splitting the time domain for the temperature series, and present the results for the subsequent fanova tests. Figures 4-16 depict the fanova procedure for all of the tests presented in this section. In them, we visualize the sample curves from the groups being tested, the bootstrap resample curves, and the empirical density under the null hypothesis. The bootstrap resample curves visualize the significance of differences between groups over time. When sample mean curves are located near the edge of the grey band formed by the resample curves, they indicate significant differences between the groups. As the sample mean curves move closer to the inside of the grey region, less significance is implied by the plot.

## 4.1 Moisture

We conduct 3 different tests on the moisture sample curves based on the functional anova procedure. Particularly, we test for differences between the 4 treatment groups, differences due to snow removal, and differences due to heating.

For testing difference between the 4 treatment groups (Figure 4), the null hypothesis is given by

$$H_0 : \mu_C = \mu_H = \mu_{HSR} = \mu_{SR} \tag{9}$$

where $\mu_C$, $\mu_H$, $\mu_{HSR}$, $\mu_{SR}$ are the expected values of the $L_2$-processes generating the control (C), heating (H), heating + snow removal (HSR), and snow removal (SR) groups respectively. The test statistic, $T_n = 0.513$, corresponds to a $p$-value of 0.074 and provides moderate evidence against $H_0$.

For testing difference due to snow removal (Figure 5), the null hypothesis is given by

$$H_0 : \mu_{\tilde{\mathrm{S}}} = \mu_{N\tilde{\mathrm{S}}} \tag{10}$$

where $\mu_{\tilde{\mathrm{S}}}$ and $\mu_{N\tilde{\mathrm{S}}}$ are the expected values of the $L_2$-processes generating the combined snow removal ($\tilde{\mathrm{S}}$) and combined no snow removal ($\mathrm{N}\tilde{\mathrm{S}}$) groups. The test statistic, $T_n = 0.097$, corresponds to a $p$-value of 0.007 and provides strong evidence against $H_0$. There is significant evidence that differences due to snow removal exist in the moisture time series. The resample curves shown in Figure 5 confirm the significance of our test. The sample mean curves are near the edge of the grey band created by the bootstrap resample curves, indicating strong evidence against the null hypothesis.

To identify difference due to heating (Figure 6), the null hypothesis is given by

$$H_0 : \mu_{\tilde{\mathrm{H}}} = \mu_{N\tilde{\mathrm{H}}} \tag{11}$$

where $\mu_{\tilde{\mathrm{H}}}$ and $\mu_{N\tilde{\mathrm{H}}}$ are the expected values of the $L_2$-processes generating the combined heating ($\tilde{\mathrm{H}}$) and combined no heating ($\mathrm{N}\tilde{\mathrm{H}}$) groups. The test statistic, $T_n = 0.029$, corresponds to a $p$-value of 0.288. There is no significant evidence of a difference due to heating in the moisture time series. This conclusion is confirmed by the resample curves shown in Figure 6. The sample mean curves are located well inside the grey band created by the bootstrap resample curves, indicating no evidence against the null hypothesis.

## 4.2 Temperature

We use the functional anova procedure given in Section 3.3 to test for differences between the 4 treatment groups, differences due to snow removal, and differences due to heating for temperature.

For testing difference between the 4 treatment groups (Figure 7, the null hypothesis is given by

$$H_0 : \mu_C = \mu_H = \mu_{HSR} = \mu_{SR} \tag{12}$$

where $\mu_C$, $\mu_H$, $\mu_{HSR}$, $\mu_{SR}$ are the expected values of the $L_2$-processes generating the control (C), heating (H), heating + snow removal (HSR), and snow removal (SR) groups respectively. The test statistic, $T_n = 542.731$, corresponds to a $p$-value of 0.150 and provides weak evidence against $H_0$.

For testing difference due to snow removal (Figure 8), the null hypothesis is given by

$$H_0 : \mu_{\tilde{\text{S}}} = \mu_{N\tilde{\text{S}}} \tag{13}$$

where $\mu_{\tilde{\text{S}}}$ and $\mu_{N\tilde{\text{S}}}$ are the expected values of the $L_2$-processes generating the combined snow removal ($\tilde{\text{S}}$) and combined no snow removal ($\tilde{\text{NS}}$) groups respectively. The test statistic, $T_n = 78.092$, corresponds to a $p$-value of 0.064 and provides moderate evidence against $H_0$.

For testing difference due to heating (Figure 9), the null hypothesis is given by

$$H_0 : \mu_{\tilde{\text{H}}} = \mu_{N\tilde{\text{H}}} \tag{14}$$

where $\mu_{\tilde{\text{H}}}$ and $\mu_{N\tilde{\text{H}}}$ are the expected values of the $L_2$-processes generating the combined heating ($\tilde{\text{H}}$) and combined no heating ($\tilde{\text{NH}}$) groups respectively. The test statistic, $T_n = 50.991$, corresponds to a $p$-value of 0.168 and provides weak evidence against $H_0$.

## 4.3 Temperature: Domain Split

A visually prominent moisture event in August corresponds to some interesting changes in the trajectories of the temperature mean curves. These mean curves for soil moisture and temperature are presented in Figure 10, and the approximate end of the large moisture event on August 24 is denoted by the dotted line. Prior to the end of the large moisture event, the mean curves from snow removal and heating + snow removal treatments correspond to higher temperatures while the mean curves from heating and control treatments correspond to lower temperatures. However, after the end of the large moisture event, the mean curves from heating and heating + snow removal treatments correspond to higher temperatures while the mean curves from snow removal and control treatments correspond to lower temperatures. It is possible that the large moisture event corresponds to some changes in the processes from which the sample curves were generated, and we believe there is justification to split the domain into two parts and consider separate functional anova tests for each part. The domain is split on August 24, the approximate end of the large moisture event, and fanova tests were conducted separately on both parts of the domain to determine whether there are difference between the 4 treatment group, difference due to Snow Removal, and difference due to Heating. The first part of the domain contains significant differences due to snow removal but not heating, and the second part of the domain contains significant differences due to heating but not snow removal. Our findings suggest that the large moisture event may have corresponded to changes in the processes from which the sample curves were generated. Detailed results for each test are provided below.

#### 4.3.1 May, 28 - Aug, 24

In this part, we consider fanova tests on temperature curves prior to August 24, the end of the large moisture event. The null hypothesis for testing differences between the 4 treatments (Figure 11) is given by

$$H_0 : \mu_C = \mu_H = \mu_{HSR} = \mu_{SR} \tag{15}$$

where $\mu_C$, $\mu_H$, $\mu_{HSR}$, $\mu_{SR}$ are the expected values of the $L_2$-processes generating the sample curves in control (C), heating (H), heating + snow removal (HSR), and snow removal (SR) groups respectively. The test statistic, $T_n = 328.627$, corresponds to a $p$-value of 0.160 and provides weak evidence against $H_0$

For testing difference due to snow removal (Figure 12), the null hypothesis is given by

$$H_0 : \mu_{\tilde{\mathrm{S}}} = \mu_{N\tilde{\mathrm{S}}} \tag{16}$$

where $\mu_{\tilde{\mathrm{S}}}$ and $\mu_{N\tilde{\mathrm{S}}}$ are the expected values of the $L_2$-processes generating the combined snow removal ($\tilde{\mathrm{S}}$) and combined no snow removal ($\mathrm{N}\tilde{\mathrm{S}}$) groups respectively. The test statistic, $T_n = 62.240$, corresponds to a $p$-value of 0.024 and provides strong evidence against $H_0$. There is significant evidence for difference due to snow removal prior to the end of the large moisture event.

For testing difference due to heating (Figure 13), the null hypothesis is given by

$$H_0 : \mu_{\tilde{\mathrm{H}}} = \mu_{N\tilde{\mathrm{H}}} \tag{17}$$

where $\mu_{\tilde{\mathrm{H}}}$ and $\mu_{N\tilde{\mathrm{H}}}$ are the expected values of the $L_2$-processes generating the combined heating ($\tilde{\mathrm{H}}$) and combined no heating ($\mathrm{N}\tilde{\mathrm{H}}$) groups respectively. The test statistic, $T_n = 9.747$, corresponds to a $p$-value of 0.549. There is no evidence of difference due to heating prior to the end of the large moisture event.

### 4.3.2 Aug, 24 - Sept 29

In this part, we consider temperature measurements after August 24, the end of the large moisture event. The null hypothesis for testing differences between the 4 treatments (Figure 14) is given by

$$H_0 : \mu_C = \mu_H = \mu_{HSR} = \mu_{SR} \tag{18}$$

where $\mu_C$, $\mu_H$, $\mu_{HSR}$, $\mu_{SR}$ are the expected values of the $L_2$-processes generating the sample curves in control (C), heating (H), heating + snow removal (HSR), and snow removal (SR) groups respectively. The test statistic, $T_n = 209.134$, corresponds to a $p$-value of 0.106 and provides moderate evidence against $H_0$.

For testing difference due to snow removal (Figure 15), the null hypothesis is given by

$$H_0 : \mu_{\tilde{\mathrm{S}}} = \mu_{N\tilde{\mathrm{S}}} \tag{19}$$

where $\mu_{\tilde{\mathrm{S}}}$ and $\mu_{N\tilde{\mathrm{S}}}$ are the expected values of the $L_2$-processes generating the combined snow removal ($\tilde{\mathrm{S}}$) and combined no snow removal ($\mathrm{N}\tilde{\mathrm{S}}$) groups respectively. The test statistic, $T_n = 9.968$, corresponds to a $p$-value of 0.336. There is no evidence of difference due to snow removal after the end of the large moisture event.

For testing difference due to heating (Figure 16), the null hypothesis is given by

$$H_0 : \mu_{\tilde{\mathrm{H}}} = \mu_{N\tilde{\mathrm{H}}} \tag{20}$$

where $\mu_{\tilde{\mathrm{H}}}$ and $\mu_{N\tilde{\mathrm{H}}}$ are the expected values of the $L_2$-processes generating the combined heating ($\tilde{\mathrm{H}}$) and combined no heating ($\mathrm{N}\tilde{\mathrm{H}}$) groups respectively. The test statistic, $T_n = 40.860$, corresponds to a $p$-value of 0.018. There is significant evidence of difference due to heating after the end of the large moisture event.

# 5 Simulation Study

The functional anova procedure assumes that the number of sample curves ($n_i$) is "large enough" for the asymptotic results given in Section 3.2 to follow. An important issue that needs to be addressed in our application of functional anova is the potential loss of power for small $n_i$. In this regard, a simulation study is conducted using sample curves with features similar to the moisture and temperature data. The goal of this study is to investigate the effect of the number of sample curves on the statistical power of the test. For simplicity, we consider functional anova tests between only 2 groups of curves. Let $\mu_1$ and $\mu_2$ be the expected values of the $L_2$-processes generating sample curves for group 1 and group 2 respectively. If $\mu_1$ and $\mu_2$ differ, we are interested in whether the functional anova procedure can correctly reject the null hypothesis, $H_0 : \mu_1 = \mu_2$.

In order to recreate features similar to our data sets, $\mu_i(t)$ and $K_i(s,t)$ are constructed from the moisture and temperature smoothed curves for the 4 scenarios.

1. In the first scenario, moisture sample curves are recreated such that $\mu_1$ and $\mu_2$ have a small difference. Let $\mu_1$ and $\mu_2$ be the mean of the

combined heating group and combined no heating group respectively from our data set. Then, $K_1(s,t)$ and $K_2(s,t)$ are the covariances of the combined heating group and combined no heating group respectively.

2. In the second scenario, moisture sample curves are recreated with a more pronounced difference between $\mu_1$ and $\mu_2$. Let $\mu_1$ be the mean of the combined heating group with a 0.02 vertical shift, and $\mu_2$ be the mean of the combined no heating group. Let $K_1(s,t)$ and $K_2(s,t)$ be the covariances of the combined heating group and combined no heating group respectively.

3. In the third scenario, temperature sample curves are recreated such that $\mu_1$ and $\mu_2$ have a small difference. Let $\mu_1$ and $\mu_2$ be the mean of the combined snow removal group and combined no snow removal group respectively from our data set. Let $K_1(s,t)$ and $K_2(s,t)$ be the covariances of the combined snow removal group and combined no snow removal group respectively.

4. In the fourth scenario, temperature sample curves are recreated such that $\mu_1$ and $\mu_2$ have a more pronounced difference. Let $\mu_1$ be the mean of the combined snow removal group with a 1 unit vertical shift, and $\mu_2$ be the mean of the combined no snow removal group. Let $K_1(s,t)$ and $K_2(s,t)$ be the covariances of the combined snow removal group and combined no snow removal group as previously defined.

The $\mu_i(t)$ and sample curves used to construct $K_i(s,t)$ for each scenario are shown in Figure 17. For each scenario and $n_i = 6, 12, 18, 24, 30$, $n_i$ sample curves are generated from a Gaussian distribution with mean $\mu_i$ and covariance $K_i(s,t)$. The functional anova procedure given in Section 3.3 is applied and repeated for 10,000 simulations. The proportion of the 10,000 simulations for which a $p$-value of less than 0.05 is obtained is given in Table 1.

For all of the scenarios, the power of the test increases as the number of sample curves $n_i$ increases. When the difference between $\mu_1(t)$ and $\mu_2(t)$ is large (scenario 2 and 4), small $n_i$ are sufficient to identify the difference. However, when the difference between $\mu_1(t)$ and $\mu_2(t)$ is small (scenario 1 and 3), larger $n_i$ are needed to identify the difference. An interesting distinction in the simulation results between the scenarios is that, overall, the temperature curves (scenario 3 and 4) have larger proportions of simulated tests with significant $p$-values compared to the moisture curves (scenario 1 and 2). This may be due to the larger covariance present in the moisture curves when compared to the temperature curves. Our simulation results indicate reasonable power for large and moderate number of sample curves. However, when the number of sample curves are small, the fanova tests lack power in identifying small differences. Additionally, even larger $n_i$ may be necessary to detect small difference if the covariance structure of the groups are large.

# 6 Conclusion

By utilizing the functional nature of our data and applying functional anova tests, we were able to identify effects due to simulated climate change on soil ecology. Particularly, significant differences were found in soil moisture due to snow removal, and evidence suggests that the processes generating the temperature curves may have changed due to a large moisture event that occurred during course of the experiment. Prior to the cessation of the large moisture event, significant difference in soil temperature are evident due to snow removal but not heating. However, after the large moisture event, significant differences in soil temperature are evident due to heating but not snow removal.

The functional anova approach is particularly useful in our application because it allows for identification of differences between groups of time series, and utilizes the behavior of the entire series. Additionally, the asymptotic fanova procedure allows us to forgo the usual one-way anova assumption of equal covariance, and the bootstrap procedure provides a computationally simple way to obtain $p$-values from which conclusions can be made. Since our approach relied on asymptotic results, the small number of sample curves per group may be cause for concern. Additionally, our analysis did not take into account any spatial dependence, which would be interesting to consider for future work.

# 7 Acknowledgements

We thank Diane Debinski from whom we obtained the data as a part of her extensive research studying the effects of simulated climate change in the montane meadows. We also thank Yehua Li for the constructive discussions regarding functional data and the smoothing procedure.

# 8 Tables

Table 1: Simulation Results giving the proportion of significant functional anova tests at $\alpha = 0.05$ for 10,000 simulations. $n_i$ denotes the number of sample curves used in the fanova procedure

| Scenario \ $n_i$ | 6 | 12 | 18 | 24 | 30 |
|---|---|---|---|---|---|
| 1 (Moist. Small Diff.) | 0.235 | 0.385 | 0.550 | 0.691 | 0.817 |
| 2 (Moist. Large Diff.) | 0.724 | 0.958 | 0.996 | 1.000 | 1.000 |
| 3 (Temp. Small Diff) | 0.539 | 0.879 | 0.984 | 0.999 | 1.000 |
| 4 (Temp. Large Diff) | 0.993 | 1.000 | 1.000 | 1.000 | 1.000 |

# 9 Figures and Tables

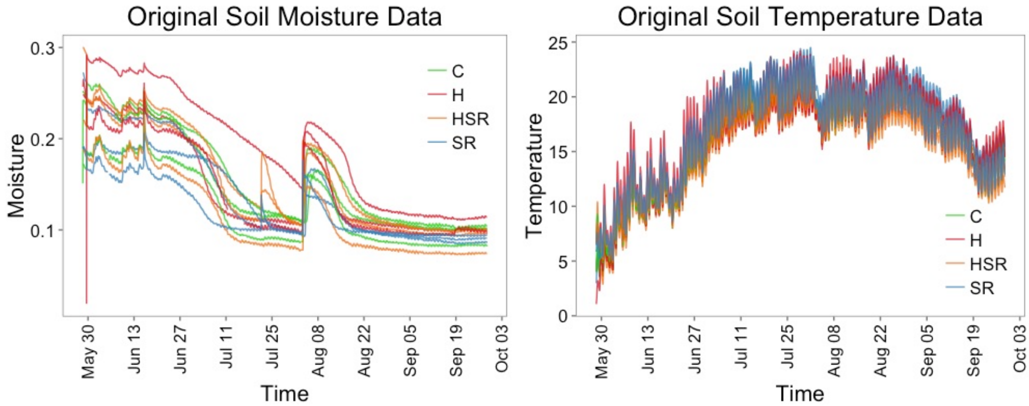


Figure 1: Left: 12 time series giving soil moisture by treatment. Right: 12 time series giving soil temperature by treatment. Treatments applied are control (C), heating (H), heating + snow removal (HSR), and snow removal (SR).

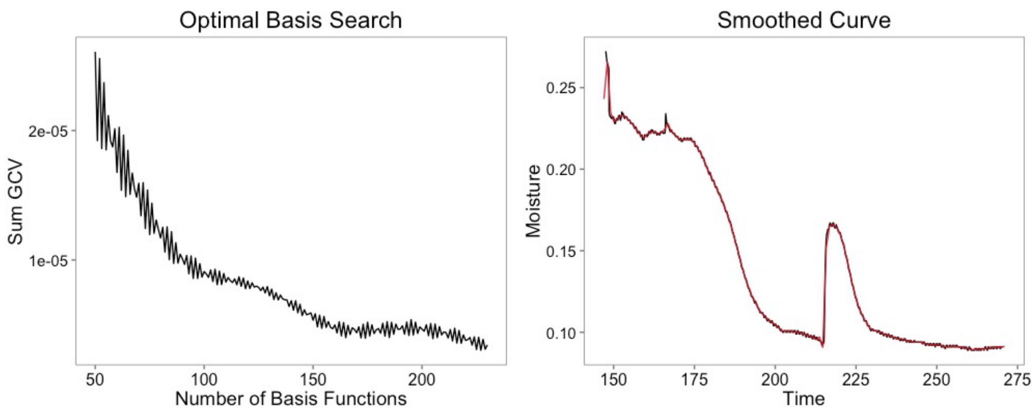


Figure 2: Left: Example plot used to obtain optimal number of basis functions in smoothing procedure. Sum of GCV criterion for moisture curves in heating treatment are plotted against number of b-spline basis functions used in smoothing. Basis number of 120 is chosen to avoid over-fitting data. Right: Plot of one functional curve (red) from heating treatment smoothed using 120 basis functions of $3^{rd}$ order overlaid on original time series (black).

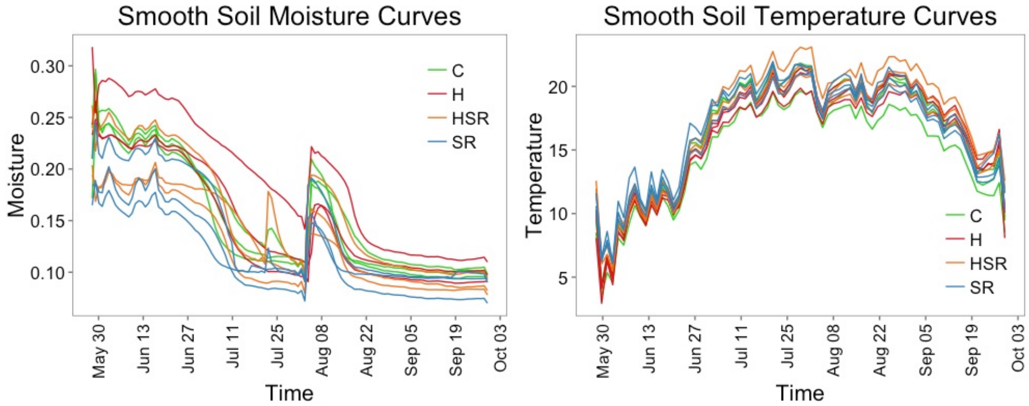


Figure 3: Left: Functional moisture curves smoothed using $3^{rd}$ order b-spline basis containing 120 elements. Right: Functional temperature curves smoothed using fourier basis containing 75 elements. Treatments applied are control (C), heating (H), heating + snow removal (HSR), and snow removal (SR).

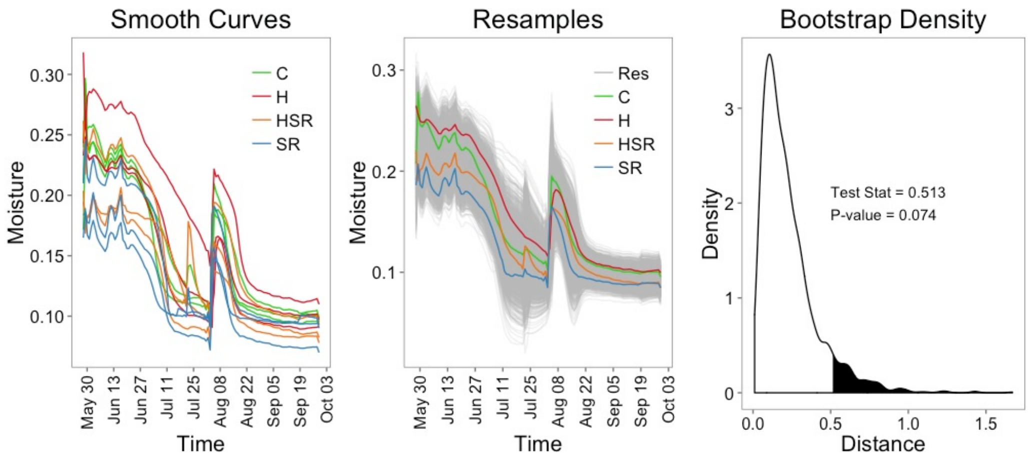


Figure 4: Fanova test for identifying differences due to the treatments in the moisture series. Left: Smooth Curves plotted by group, where the groups are control (C), heating (H), heating + snow removal (HSR), and snow removal (SR). Middle: Bootstrap resample curves shown in grey, and mean sample curves shown in color. Right: Bootstrap density with the shaded region indicating the $p$-value.

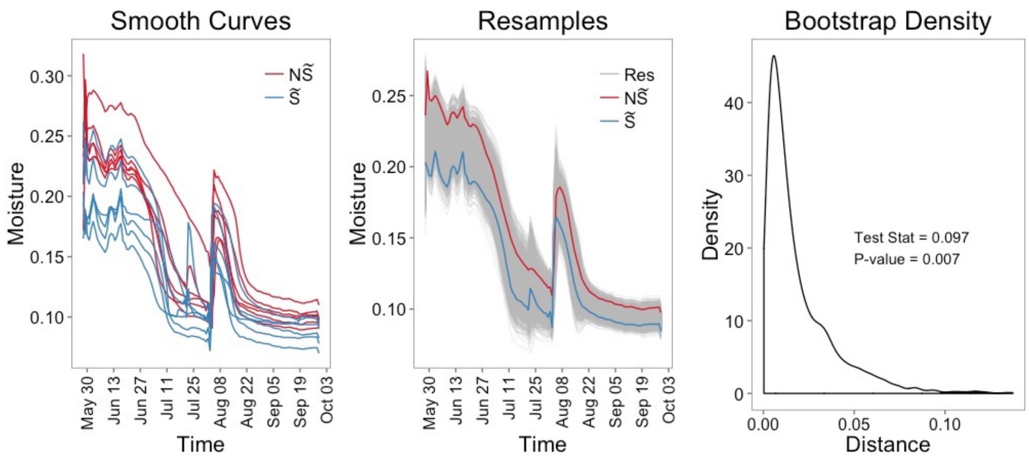


Figure 5: Fanova test for identifying differences due to snow removal in the moisture series. Left: Smooth Curves plotted by group, where the groups are combined snow removal (S̃) and combined no snow removal (NS̃). Middle: Bootstrap resample curves shown in grey, and mean sample curves shown in color. Right: Bootstrap density with the shaded region indicating the $p$-value..

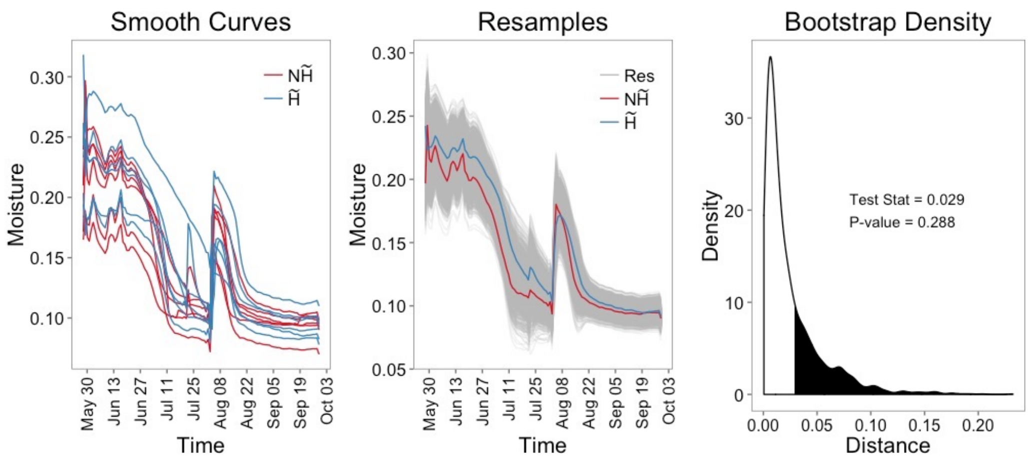


Figure 6: Fanova test for identifying differences due to heating in the moisture series. Left: Smooth Curves plotted by group, where the groups are combined heating (H̃) and combined no heating (NH̃). Middle: Bootstrap resample curves shown in grey, and mean sample curves shown in color. Right: Bootstrap density with the shaded region indicating the $p$-value.

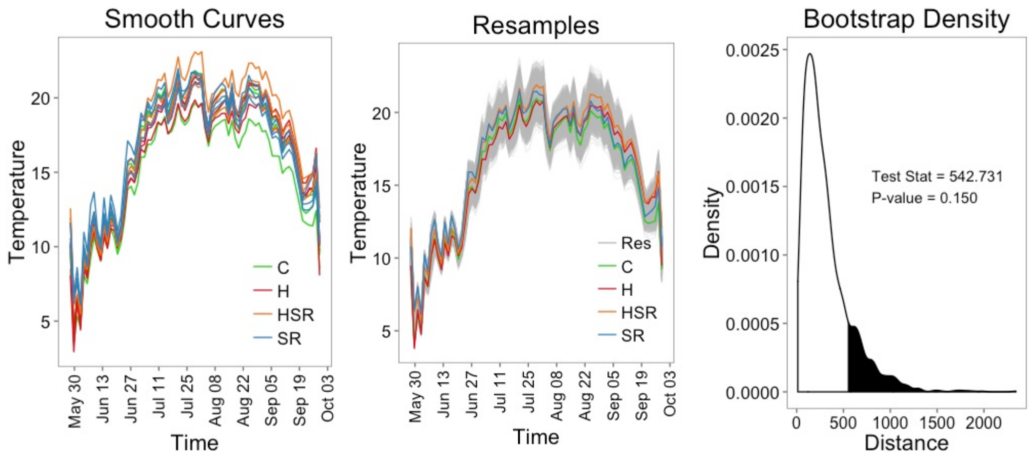


Figure 7: Fanova test for identifying differences due to the treatments in the temperature series. Left: Smooth Curves plotted by group, where the groups are control (C), heating (H), heating + snow removal (HSR), and snow removal (SR). Middle: Bootstrap resample curves shown in grey, and mean sample curves shown in color. Right: Bootstrap density with the shaded region indicating the $p$-value.

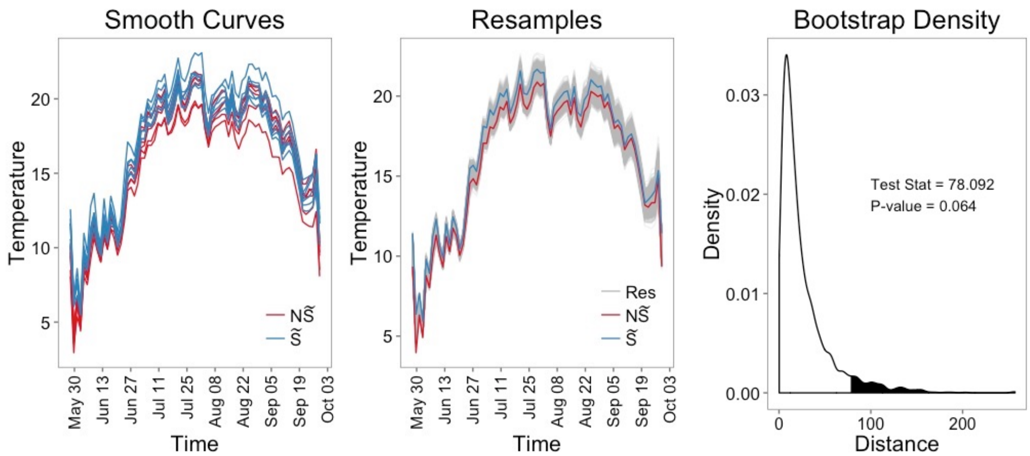


Figure 8: Fanova test for identifying differences due to snow removal in the temperature series. Left: Smooth Curves plotted by group, where the groups are combined snow removal ($\tilde{S}$) and combined no snow removal ($N\tilde{S}$). Middle: Bootstrap resample curves shown in grey, and mean sample curves shown in color. Right: Bootstrap density with the shaded region indicating the $p$-value.

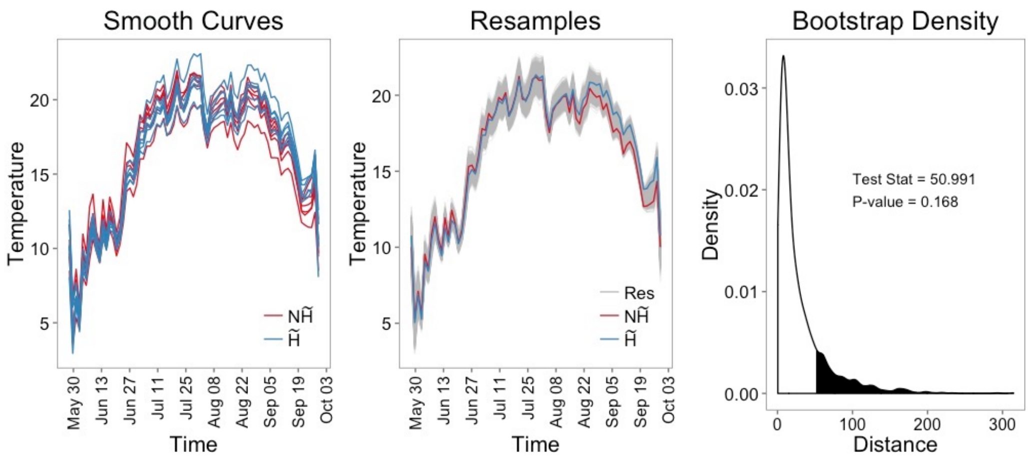


Figure 9: Fanova test for identifying differences due to heating in the temperature series. Left: Smooth Curves plotted by group, where the groups are combined heating ($\tilde{\mathrm{H}}$) and combined no heating (N$\tilde{\mathrm{H}}$). Middle: Bootstrap resample curves shown in grey, and mean sample curves shown in color. Right: Bootstrap density with the shaded region indicating the $p$-value.

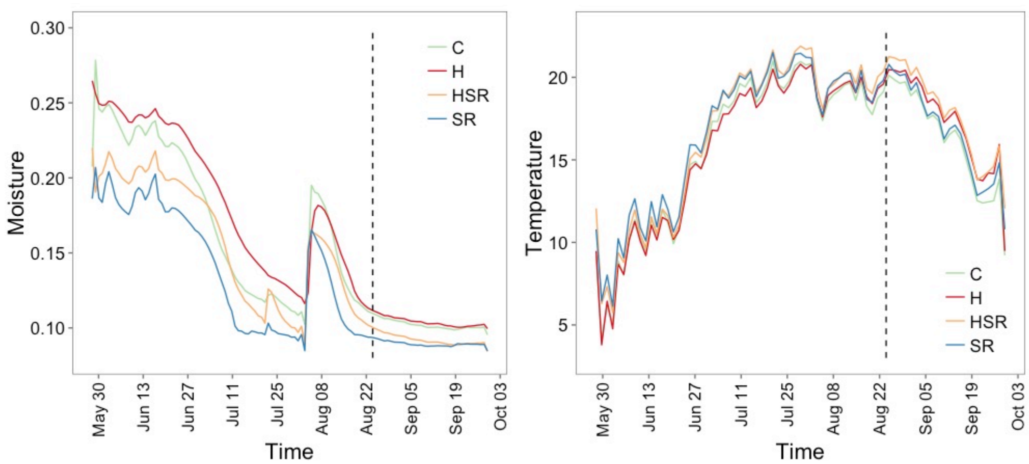


Figure 10: Left: Smooth moisture series showing a large moisture event and dotted line indicating the approximate end of the event, August 24. Right: Smooth temperature series with the dotted line indicating the end of the large moisture event. The domain for temperature will be split for further analysis at the dotted line, and fanova tests are conducted separately on each part of the domain.

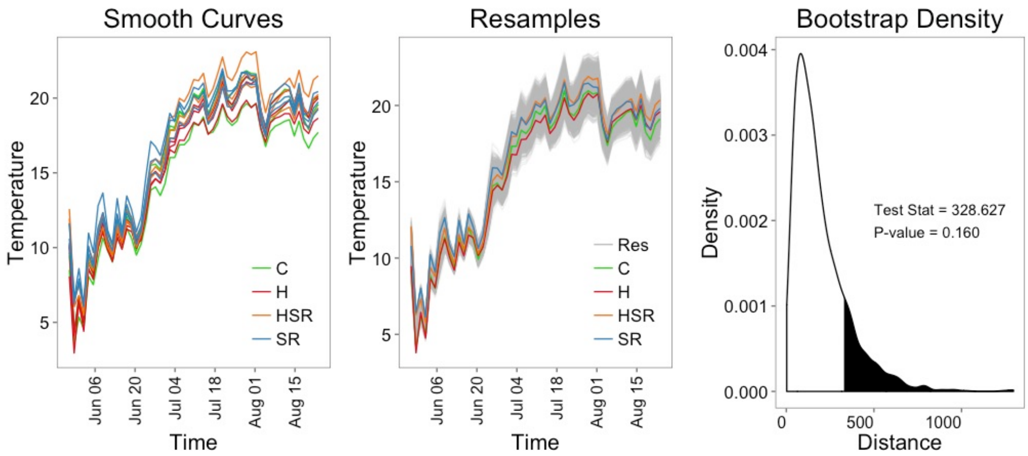


Figure 11: Fanova test for identifying differences due to the treatments in the temperature series between May 30, 2011 and August 24, 2011. Left: Smooth Curves plotted by group, where the groups are control (C), heating (H), heating + snow removal (HSR), and snow removal (SR). Middle: Bootstrap resample curves shown in grey, and mean sample curves shown in color. Right: Bootstrap density with the shaded region indicating the $p$-value.

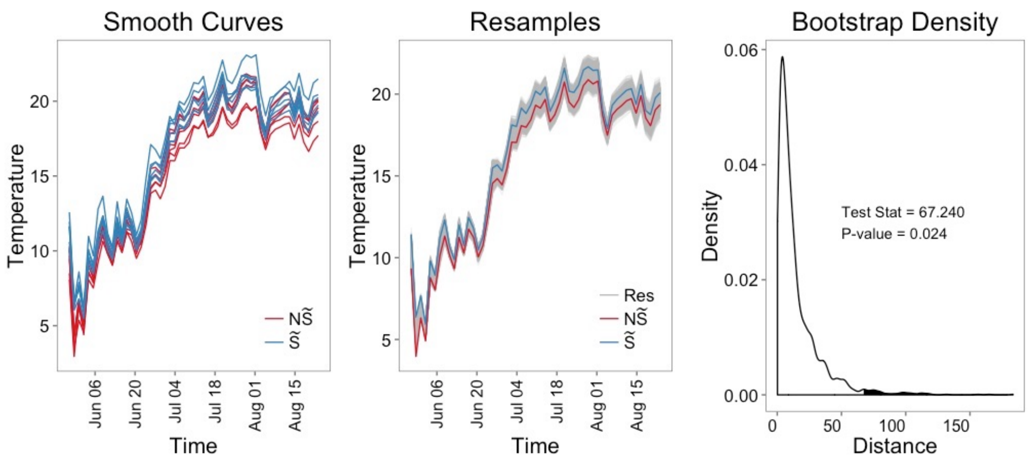


Figure 12: Fanova test for identifying differences due to snow removal in the temperature series between May 30, 2011 and August 24, 2011. Left: Smooth Curves plotted by group, where the groups are combined snow removal ($\tilde{\mathrm{S}}$), and combined no snow removal (N$\tilde{\mathrm{S}}$). Middle: Bootstrap resample curves shown in grey, and mean sample curves shown in color. Right: Bootstrap density with the shaded region indicating the $p$-value.

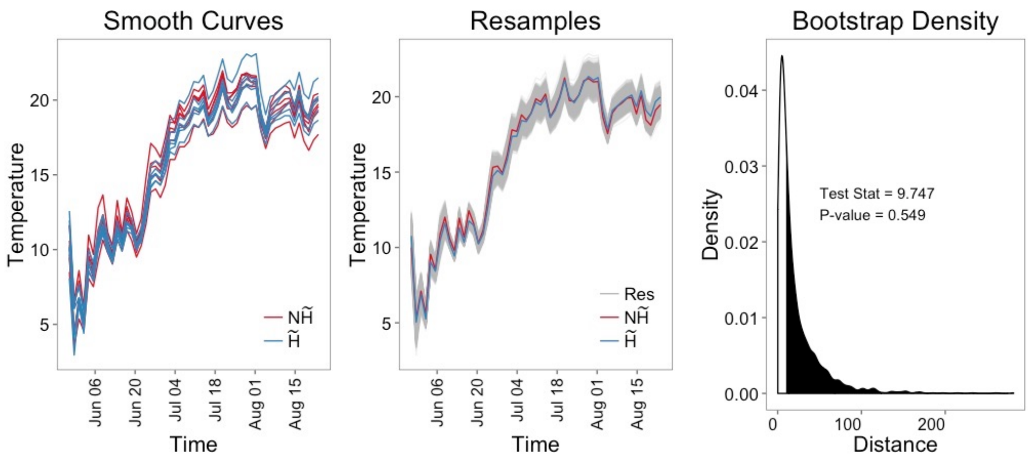


Figure 13: Fanova test for identifying differences due to heating in the temperature series between May 30, 2011 and August 24, 2011. Left: Smooth Curves plotted by group, where the groups are combined heating ($\tilde{\text{H}}$), and combined no heating ($\text{N}\tilde{\text{H}}$). Middle: Bootstrap resample curves shown in grey, and mean sample curves shown in color. Right: Bootstrap density with the shaded region indicating the $p$-value.

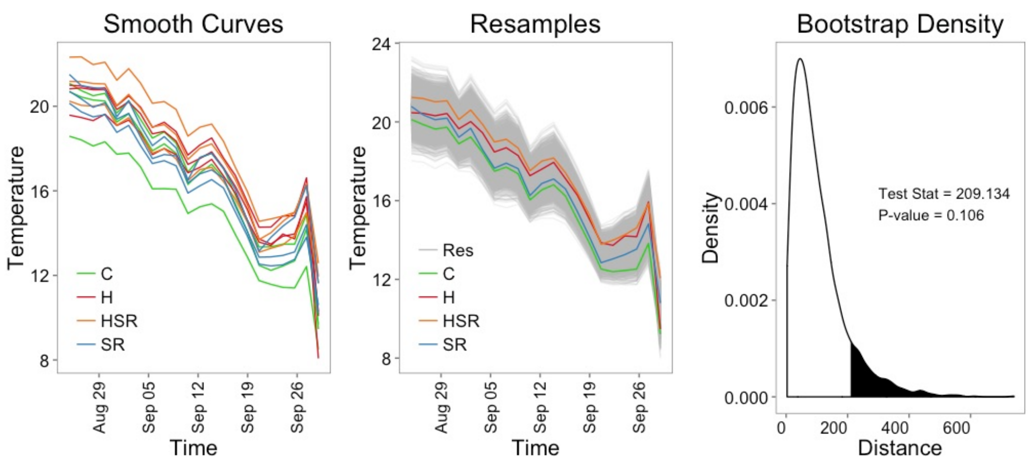


Figure 14: Fanova test for identifying differences due to the treatments in the temperature series between August 24, 2011 and September 29, 2011. Left: Smooth Curves plotted by group, where the groups are control (C), heating (H), heating + snow removal (HSR), and snow removal (SR). Middle: Bootstrap resample curves shown in grey, and mean sample curves shown in color. Right: Bootstrap density with the shaded region indicating the $p$-value.

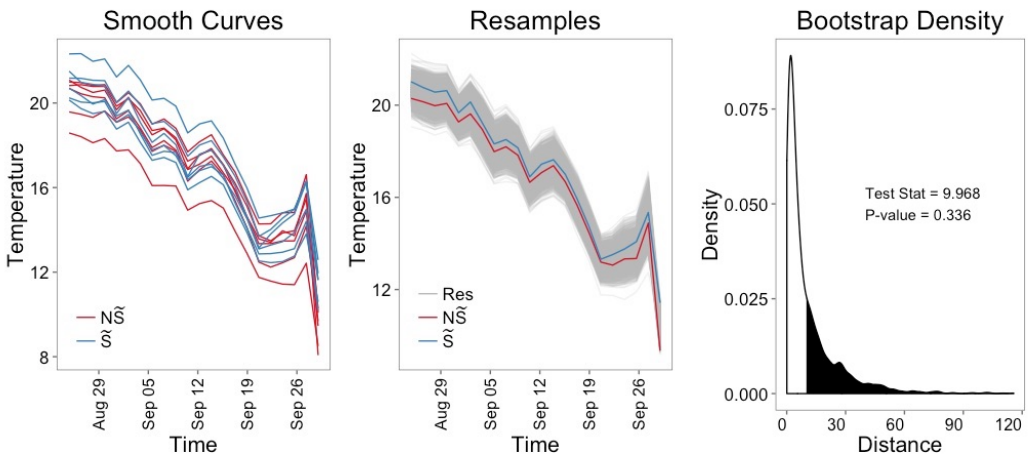


Figure 15: Fanova test for identifying differences due to snow removal in the temperature series between August 24, 2011 and September 29, 2011. Left: Smooth Curves plotted by group, where the groups are combined snow removal ($\tilde{S}$), and combined no snow removal ($N\tilde{S}$). Middle: Bootstrap resample curves shown in grey, and mean sample curves shown in color. Right: Bootstrap density with the shaded region indicating the $p$-value.

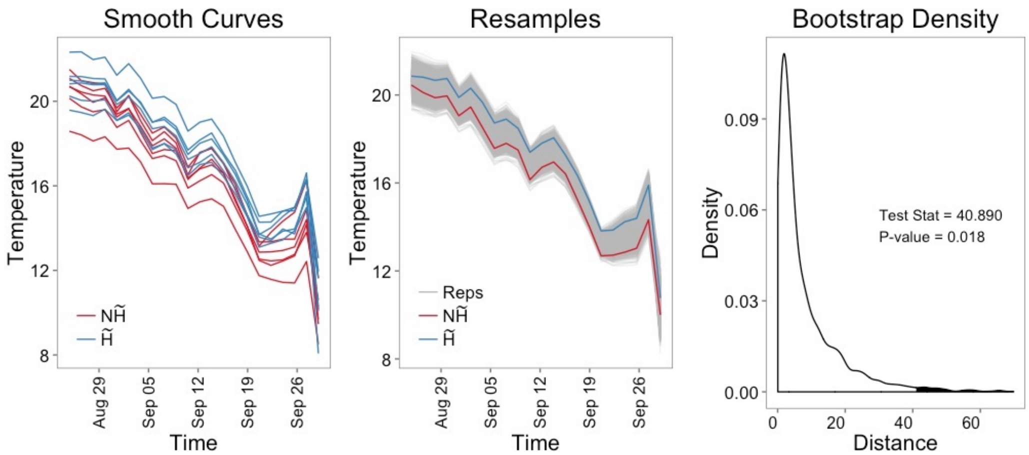


Figure 16: Fanova test for identifying differences due to heating in the temperature series between August 24, 2011 and September 29, 2011. Left: Smooth Curves plotted by group, where the groups are combined heating ($\tilde{H}$), and combined no heating ($N\tilde{H}$). Middle: Bootstrap resample curves shown in grey, and mean sample curves shown in color. Right: Bootstrap density with the shaded region indicating the $p$-value.

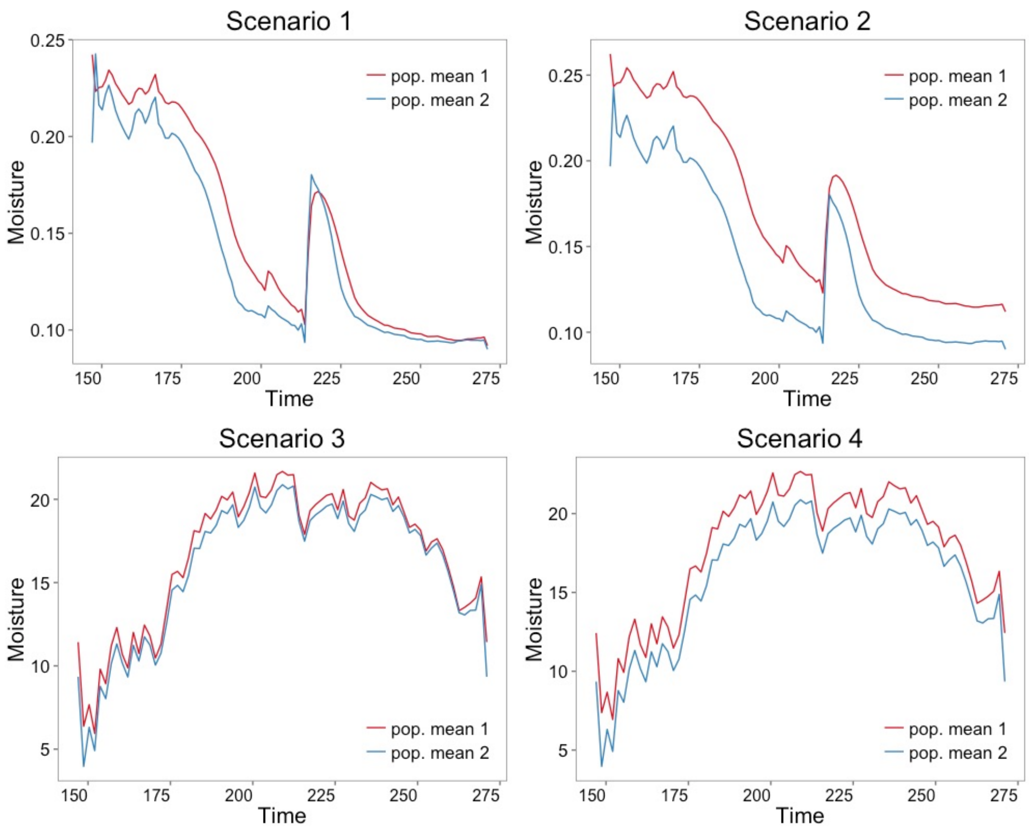


Figure 17: Population mean curves from which sample curves in simulation study were generated are shown. Top Left: Scenario 1 showing small difference between curves of 2 groups with features resembling moisture data. Top Right: Scenario 2 showing large difference between 2 groups with features resembling moisture data. Bottom Left: Scenario 3 showing small difference between 2 groups with features resembling temperature data. Bottom Right: Scenario 4 snowing large difference between 2 groups with features resembling temperature data.

# References

[1] A. T. Classen, M. K. Sundqvist, JA Henning, G. S. Newman, J. A. M. Moore, M. A. Cregger, L.C. Moorhead, and C.M. Patterson. Direct and indirect effects of climate change on soil microbial and soil microbial-plant interactions: What lies ahead? *Ecosphere*, 6(8):130, 2015.

[2] A. Cuevas, M. Febrero, and R. Fraiman. An anova test for functional data. *Computational Statistics and Data Analysis*, 47(1):111–122, 2004.

[3] J. Ramsay and B. Silverman. *Functional Data Analysis*. Springer, 2 edition, 2005.

[4] J. A. Sherwood, D. M. Debinski, P. C. Caragea, and M. J. Germino. Effects of experimentally reduced snowpack and passive warming on montane meadow plant phenology and floral resources. *Ecosphere*, 8(3), 2017.

[5] J. Tarrio-Saaverda, S. Naya, M. Francisco-Fernandez, R. Artiaga, and J. Lopez-Beceiro. Application of functional anova to the study of thermal stability of micronano silica epoxy composites. *Chemometrics and Intelligent Laborotory Systems*, 105:114–124, 2011.

[6] G. Wahba and P. Craven. Smoothing noisy data with spline functions. estimating the correct degree of smoothing by the method of generalized cross-validation. *Numerische Mathematatik*, 31:377–404, 1978/79.