Pollinator Diversity in the Niwot Ridge LTER Black Sands Experiment

A Causal Reanalysis by Treson Thompson

Introduction

Background

Alpine plant communities are fragile ecosystems that provide important ecosystem services and have a cascade effect on regions further down the watershed. One vital component of these communites are the pollinators that promote genetic diversity and allow for propogation.

The Niwot Ridge LTER Black Sands Experiment was established in 2018 to study the effects of early snow melt on these fragile plant networks. Black sand was scattered on the surface of the snow in early spring to warm up in the sun and melt the snow prematurely while minimizing the unwanted effects of other interventions. To control for the addition of sand to the environment, an equal amount was scattered on control plots after all the snow had melted naturally. This sand was applied seasonally from 2018 to 2024. Five experimental sites were established non-randomly to capture landscape heterogeneity (Soddie, Trough, Audobon, Lefty, and East Knoll). Each year, data were collected on snow depth across these sites, along with data on plant community composition. Here, the effect of early snow melt on pollinator diversity in these sites is investigated, using data collected by Rose-Person et. al., 2024b on pollinator visitations in 2020 in the experimental sites. While Rose-Person et. al. found an effect of early snowmelt on some of the drivers of pollinator networks, these data are here reevaluated using causal inference methods.

Directed Acyclic Graph

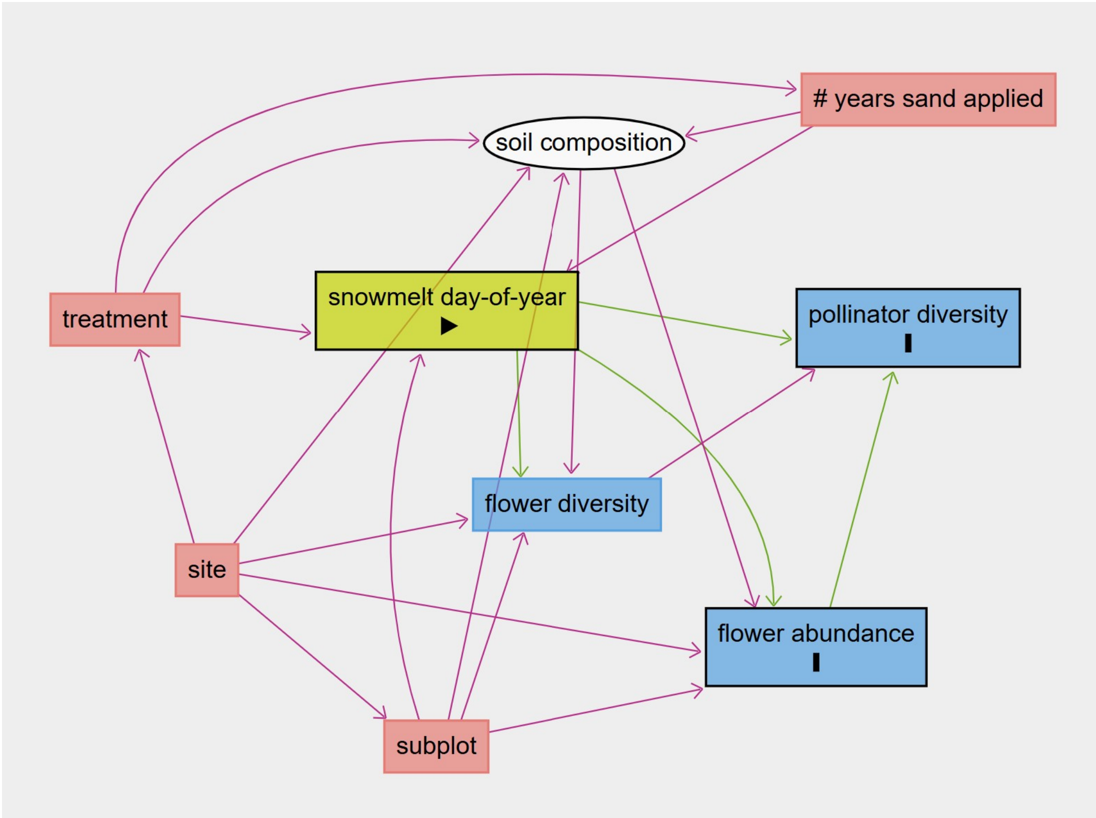

Figure 1. Directed acyclic graph showing the causal relationships between variables in this study system. Omitted variables are assumed to not have a causal impact on this system. Variables are described in Table 1.

Variable	Description	Data source	Assumptions
treatment	Binary; application of black sand before or after snow melt.	Forrester and Morse, 2024	Sand applied evenly and consistently
site	Categorical; five study sites. (Lefty excluded in Rose-Person data)	Forrester and Morse, 2024	Sites are comparable and generalizable
subplot	Categorical; five subplots within in site treatment, arranged from high to low elevation	Forrester and Morse, 2024	Subplots are comparable across treatment levels
soil composition	Unobserved		Any changes affect all sites evenly
# years sand applied	Ordinal; number of experimental seasons	Forrester and Morse, 2024	Not site-variant
snow melt day-of-year (DOY)	Various metrics	Forrester and Morse, 2024	Metrics are quantified correctly
flower diversity	Not quantified for this	Rose-Person et. al., 2024a	Does not create a back door pathway

	analysis		
flower abundance	Not quantified	Rose-Person et. al., 2024a	Does not create a back door pathway
pollinator diversity	Various metrics	Rose-Person et. al., 2024a	Metrics are quantified correctly

Table 1. Descriptions of the variables presented in Figure 1.

Methods

Data packages were acquired from the Niwot Ridge Long-term Ecological Research Station (LTER). The experiment was conducted at five separate sites on Niwot Ridge (Audubon, East Knoll, Soddie, Trough, and Lefty), their locations chosen non-randomly to capture heterogeneity in the landscape. Each site was 40 x 20 meters, with the long axis running from high to low elevation. Each site was split along the short axis into a treatment plot and a control plot, creating two 40 x 10 m plots. Forrester and Morse, 2024 provided snow depth measurements at all five experimental sites, with five measurements taken at regular intervals from high to low elevation within each treatment plot from 2018 to 2024. Snow depth was measured from early spring until the snow was gone from a particular plot to capture the snow melt process.

Data on pollination events were collected by Rose-Person et. al., 2024a in the spring and summer of 2020. Five subplots were established within each experimental plot at four of the Black Sands sites, for a total of 40 subplots. Pollinator and flower diversity and abundance were recorded at each subplot, along with the slope, aspect, elevation, and topographic position index at eleven meters. For further details on pollinator data collection, reference Rose-Person et. al., 2024b.

First, an analysis of the effect of treatment on the Simpson’s diversity index of individual subplots was conducted using a traditional linear model using observed confounders as controls. This traditional ecological model was specified in R with treatment as the predictor variable, Simpson’s diversity index (calculated from the total number of pollinators within a subplot over the course of the season) as the response variable, site as a fixed effect, and controlling for subplot slope, aspect, elevation, and TPI at an 11 m radius.

The next step in the analysis was to justify the application of black sand as an effector for early snow melt. Because pollinator diversity data was only collected during one season, the backdoor pathway created by successive application of black sand cannot be closed. However, if treatment is an appropriate instrumental variable for early snow melt, it can be assumed that variance created by successive treatment is a year-level time-variant confounder while remaining site-invariant. Application of black sand was assessed as an instrumental variable for early snow melt in two ways. First, early snow melt was quantified as the relationship between snow depth and day-of-year. A smoothed generalized additive model (GAM) was used to assess this relationship, where the predicted smooth was allowed to vary by treatment, using the “mgcv” package in R. This IV treatment GAM was then compared to a GAM where the smooth was not allowed to vary by treatment, the IV null GAM. An

ANOVA was used to make this comparison to determine if the variation in snow depth over time (day-of-year) was better explained when treatment was taken into account. Second, early snow melt was quantified as the first day-of-year on which no snow was recorded at a particular subplot. A more traditional standard two-way fixed effects linear model (TWFE LM) was used to regress snow melt day-of-year on the treatment with site, subplot, and year applied as fixed effects to control from both site-level and year-level variance and block back door pathways from “treatment” to “early snow melt” (Figure 1).

The second step of the analysis was to use treatment as an instrument to assess the causal effect of early snow melt on pollinator diversity and estimate the average treatment effect (ATE). Pollinator diversity was calculated using Simpson’s diversity index and as alpha diversity from observations of species in a particular subplot. This diversity was calculated as the peak Simpson’s and alpha diversity DOY (as in Rose-Person et. al. 2024b, n = 42), total diversity over the course of the season at each subplot (n = 42), and as the smoothed prediction of DOY on diversity (n = 508) to further investigate if treatment had an effect on diversity over time. The first metric was assessed using a linear model (Diversity LM Peak) with the DOY on which Simpson’s or alpha diversity was highest as the outcome, treatment as the predictor, and site and subplot as fixed effects. The second metric was assessed using a linear model (Diversity LM Total) with season-total diversity within a subplot as the outcome variable, treatment as the predictor, and site and subplot as fixed effects. The third metric was fitted using a GAM (Diversity GAM Treatment), allowing the smoothed spline to vary by treatment. This was compared to a null GAM (Diversity GAM Null) where all units were considered evenly. An ANOVA was then used to compare these two models and assess if including treatment in the model created a significantly better smoothed spline.

A series of robustness checks were also performed to evaluate assumptions made about the experimental design and causal inference methodology. To determine if subplots were comparable across treatment groups, a generalized linear model was used to evaluate if any site-level confounders predicted treatment assignment. To this end, a GLM was fitted with treatment as the outcome variable and slope, elevation, and TPI at 11 m as confounders. Wilcoxon tests were also run to assess balance between slope, elevation, and TPI at 11 m between control and treated subplots. Only the data from 2020 were used for this model, as this was the only year that pollinator data were collected and because these site-variant variables are assumed to not vary from year to year.

Results

The traditional model found no significant effect of snow melt DOY on pollinator diversity (p = 0.3677).

Early snow melt was first assessed as the smoothed relationship between snow depth and day-of-year using GAMs. The treatment and null models both showed a highly significant correlation between snow depth and DOY. Relevant values are recorded in Table 2. The ANOVA comparing these two models found a significant different between the two (deviance: -51272, p = 0.02447), with the model allowing the smooth to vary by treatment explaining more of the overall variance (Figure 2).

Model	Description	Predictor	Outcome	Analytical Unit	Specification	Estimate	SE	p-value
Traditional	Assess effect of treatment on Simpson's diversity index	Treatment	Simpson's diversity index	Subplot (7 years)	Simpson diversity ~ treatment + site + elevation + tpi + aspect + slope	0.0014	0.0015	0.3677
IV GAM Treatment	Predict snow depth based on DOY and treatment	Treatment	Snow depth	Subplot (7 years)	Snow depth ~ s(DOY, by = treatment) + treatment + site + subplot	F-statistic (treatment): 236.3; (control): 438.3		Treatment: 2e-16***; Control: 2e-16***
IV GAM Null	Predict snow depth based on DOY	Treatment	Snow depth	Subplot (7 years)	Snow depth ~ s(DOY) + treatment + site + subplot	474.4		2e-16***
IV Comparison ANOVA	Compare IV GAMs to assess IV validity			GAM model		Deviance: -78295		0.00123*
IV LM	Predict early snow melt based on treatment	Treatment	Snow melt DOY	Subplot (7 years)	Snow melt DOY ~ treatment + site + subplot	-6.247	1.39	1.07e-5**
Diversity GAM Treatment	Predict diversity based on DOY and treatment	Day-of-year	Simpson's diversity index	Subplot (1 year)	Simpson's diversity ~ s(DOY, by = treatment) + treatment + site + subplot	F-statistic (treatment): 2.233; (control): 3.61		Treatment: 0.053; Control: 0.0022*
Diversity GAM Null	Predict diversity based on DOY	Day-of-year	Simpson's diversity index	Subplot (1 year)	Simpson's diversity ~ s(DOY) + treatment + site + subplot	F-statistic: 4.783		2.27e-5**
Diversity Comparison ANOVA	Compare Diversity GAMs to assess if treatment alters diversity over time			GAM model		-0.1397		0.9561
Diversity LM Peak (Simpson's)	Predict diversity based on early snow melt (ATE)	Treatment	DOY on which Simpson's diversity was highest	Subplot (1 year)	Simpson highest DOY ~ treatment + site + subplot	-10.00	3.52	0.008*
Diversity LM Peak (Richness)	Predict diversity based on early snow melt (ATE)	Treatment	DOY on which alpha diversity was highest	Subplot (1 year)	alpha highest DOY ~ treatment + site + subplot	-9.38	3.63	0.014*
Diversity LM Total (Simpson's)	Predict diversity based on early snow melt (ATE)	Treatment	Simpson's diversity index	Subplot (1 year, pooled)	Simpson's diversity ~ treatment + site + subplot	-0.06	0.028	0.042*
Diversity LM Total (Richness)	Predict diversity based on early snow melt (ATE)	Treatment	Alpha diversity	Subplot (1 year, pooled)	Richness ~ treatment + site + subplot	-0.476	0.637	0.46
Diversity LM Peak LATE (Simpson's)	Predict diversity based on early snow melt (LATE)	Treatment	DOY on which Simpson's diversity was highest	Compliant subplot (1 year)	Simpson highest DOY ~ treatment + site + subplot	-13.8	5.47	0.022*
Diversity LM	Predict	Treatment	DOY on which	Compliant	alpha highest	-12.6	5.82	0.045*

Peak LATE (Richness)	diversity based on early snow melt (LATE)		alpha diversity was highest	subplot (1 year)	DOY ~ treatment + site + subplot			
Diversity LM Total LATE (Simpson's)	Predict diversity based on early snow melt (LATE)	Treatment	Simpson's diversity index	Compliant subplot (1 year, pooled)	Simpson's diversity ~ treatment + site + subplot	-0.105	0.039	0.015*
Diversity LM Total LATE (Richness)	Predict diversity based on early snow melt (LATE)	Treatment	Alpha diversity	Compliant subplot (1 year, pooled)	Richness ~ treatment + site + subplot	-0.667	0.908	0.473

Table 2. Descriptions and outputs for the models discussed in this analysis. The Estimate is the estimated beta for linear models unless otherwise specified. Asterisks indicate the level of significance.

The treatment's effect on early snow melt was also assessed using the snow melt DOY (first day of the year where the snow depth was zero) as the quantifier at the subplot level over the course of six snow melt seasons (2018 – 2024). The IV LM also found a significant relationship between the IV and the treatment variable snow melt DOY. The estimated slope was -6.2468 for the effect of treatment on snow melt DOY (SE: 1.3969, p = 1.07e-5). This model had an F-statistic of 47.24 on 10 and 331 degrees of freedom (p < 2.2e-16), further demonstrating a credible relationship between treatment and snow melt DOY.

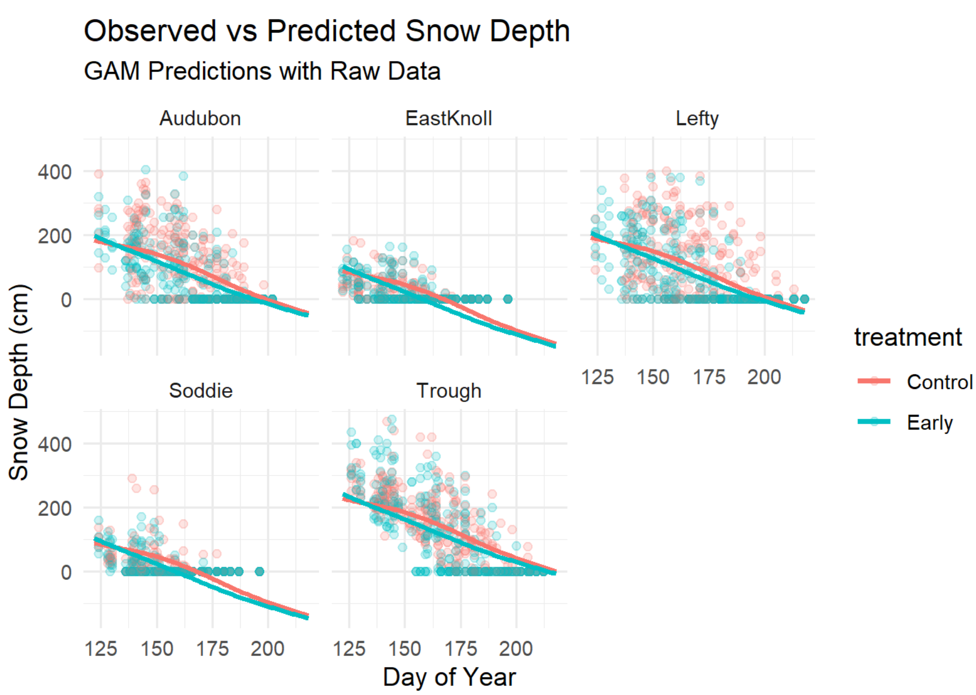

Figure 2. Smoothed splines predicting snow depth over time in five study sites. Pooled across six years of data. Points represent actual values recorded at subplots, lines represent smoothed predictions.

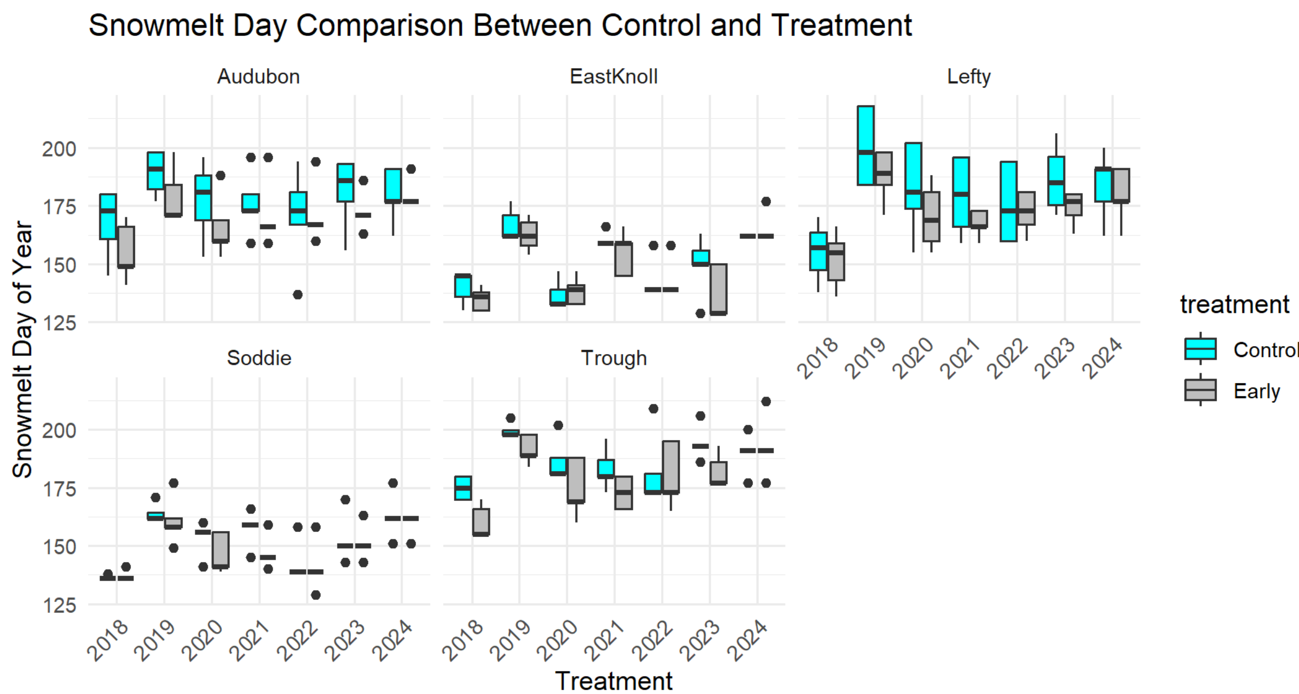

Figure 3. Average subplot snow melt DOY across the five study sites split by year. A visualization of the IV LM.

The assessment of the effect of early snow melt as effected by black sand application on pollinator diversity within subplots (ATE) first examined the DOY on which Simpson's or alpha diversity was highest within subplots. The Diversity LM Peak found a significant effect of treatment on peak Simpson's diversity DOY (Estimate = -9.381, SE = 3.628, p = 0.0145). This model also found a significant effect of treatment on peak species richness DOY (Estimate = -10.00, SE = 3.521, p = 0.0078). The second metric assessed Simpson's diversity as calculated from the summed observations within subplots as predicted by treatment. The Diversity LM Total found no effect of treatment on the Simpson's diversity index of subplots (p = 0.2615). Neither was there an effect in this model when diversity was quantified as species richness (p = 0.1798). Considering Simpson's diversity index as a function of time over the course of summer 2020, the treatment and null GAMs both found a significant relationship between the DOY and the pollinator diversity, but the ANOVA Chi-squared test comparing the two models found no significant difference between the two (df = -6.7264, deviance = -0.13971, p = 0.9561).

Additionally, attempts to estimate the LATE using only subplot pairs where the control subplot melted after the treatment subplot found similar results. For example, treatment advanced peak Simpson's and alpha diversity by 13.8 (SE = 5.47, p = 0.022) and 12.6 (SE = 5.82, p = 5.82) days respectively. Treatment also significantly decreased overall Simpson's diversity by -0.105 (SE = 0.039, p = 0.015), but did not significantly affect richness (p = 0.473). Although the Diversity LM Total LATE for richness had a negative effect of treatment on richness, this model was nearly a perfect fit, suggesting there were not enough compliant treatment units or not enough variance in snow melt DOY to make this estimate.

Robustness checks to establish even pairing of control and treatment subplots showed that covariate confounders were not significant predictors of treatment status. The GLM found no significant effect of slope, elevation, or TPI at 11 m on treatment status. The Wilcoxon tests likewise found no significant effect of any of these covariates on treatment status.

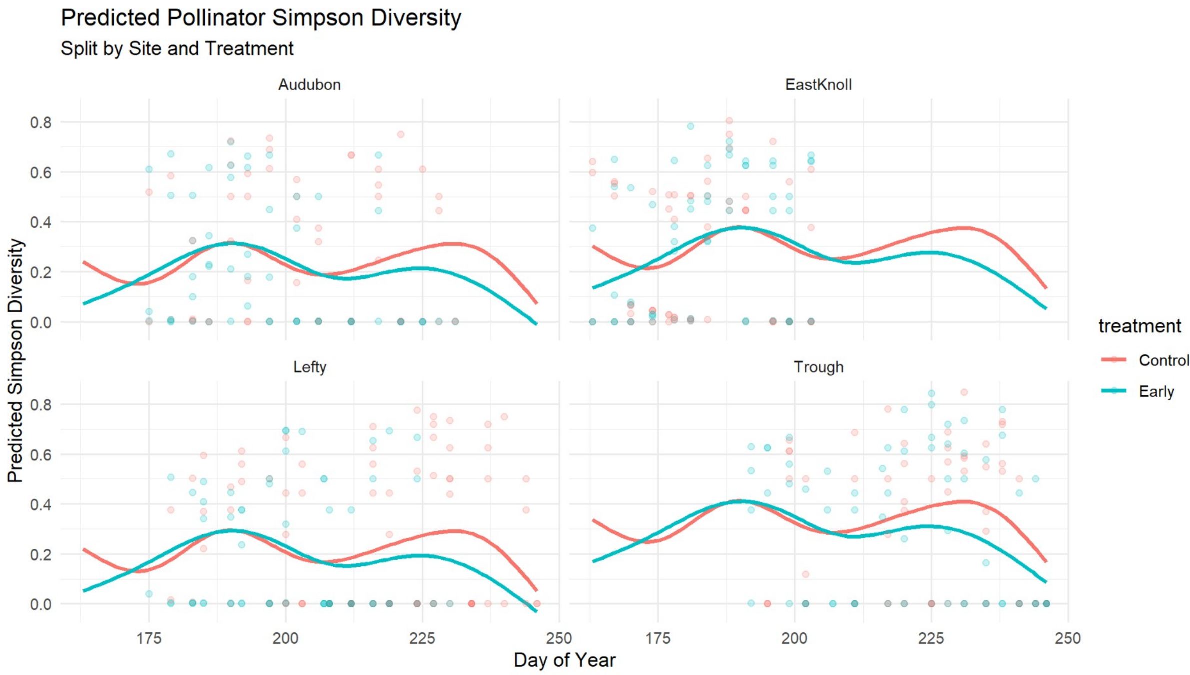

Figure 4. Smoothed splines of subplot pollinator diversity (Predicted Simpson Diversity) over time (Day of Year) in four of the experimental sites.

Discussion

In comparison to the traditional model specified using incorrect or unspecified causal assumptions, the models presented in this analysis present more consistent and significant estimates. As shown in the results section, models specified using more relaxed causal assumptions than the Traditional LM can change the over conclusions drawn from the analysis. Thus, is it important to address and specify the causal assumptions used in an analysis in order to correctly specify analytical models and draw generalizations from the estimands of interest.

Assumptions

This anlaysis makes the stable unit treatment value assumption (SUTVA) that outcomes for individual subplots are not impacted by other units and that units were treated homogeneously within treatment groups, and that measurement error was minimal across all measured variables. The assumption that the backdoor path opened by potential accumulation of black sand over multiple years of treatment can be blocked by controlling for experimental year when assessing the effect of black sand is also made. While Rose-Person et. al. made the strong assumption that all site-invariant confounding variables were measured and controlled for, here this assumption is relaxed by implementing site- and subplot-level fixed effects.

When assessing the IV, it was assumed that all confounding variables that may impact snow melt and be correlated with the error term were subsumed by our subplot, site, and year fixed effects. The site and subplot fixed effects control for site-variant confounders such as slope, elevation, and TPI at 11 m. The year fixed effect controls for time-variant confounders such as weather, annual precipitation, buildup of black sand, or wind conditions. It is also assumed that applying black sand in the control subplots after snow melt did not have any relevant effect, and if it did, that this effect would be uniform across both treatment and control plots. For example, it is assumed that sand applied on top of the snow was not removed from the plot through hydrological processes at a different rate from the sand applied directly onto the tundra, or that sand applied in the control plots did not elevate the ground temperature more than the treatment plots. Essentially, this assumption is that the control subplots serve as effective counterfactuals to the treatment subplots.

For the examination of the causal effect of early snow melt on pollinator diversity, it was assumed that treatment with black sand was an appropriate and strong instrument for early snow melt. In order for the ATE to be generalizable to high alpine ecosystems, the assumption must be made that application of black sand only affected the system through the advancement of snow melt. However, this assumption may not be valid if the quantity of sand applied to the experimental plots caused soil composition in the experimental sites to differ relevantly and significantly from the surrounding ecosystem. The estimation of the ATE assumes that treatment effectively advanced snow melt during the year in which pollinator data was collected in order to control for time-variant confounders. Since pollinator data were only collected during one season, including a fixed effect for year was not possible. Thus, the LATE may be a more appropriate estimand as the research question is interested in the effect of early snow melt on pollinator diversity, not on the effect of black sand application. Non-compliant subplots are thus biologically irrelevant.

IV Justification

Conception of the treatment as an IV for early snow melt is supported on several levels. For one, treatment significantly alters the variance in snow depth over time, as shown by the ANOVA comparing IV GAMs. These GAMs showed a highly significant correlation between snow depth and DOY, as expected, since snow depth decreases as summer approaches. The metric of interest is the ANOVA which found that including two separate smoothed splines in the model, one for each treatment, did a significantly better job of explaining the variation in snow depth over time. Secondly, the IV LM showed a significant effect of treatment on snow melt DOY when controlling for site, subplot, and year fixed effects. The F-statistic for this model was above the traditional threshold of 40 for strong IV support as well. This suggests the assumption that within the 2020 season in which pollinator data were collected, treatment can be used as an overal proxy to estimate ATE despite the 18 subplots that did not technically comply with treatment during that season.

However, these models assessed compliance with treatment on average, i.e. subplots were not paired with their control-treatment counterpart directly. The number of non-compliant treatment-control subplot pairs within each year suggests that black sand application may not be a *reliable* effector for early snowmelt, although it does seem to advance snow melt on average. For the purposes of this analysis, however, the quantitative support for the IV is heeded.

Estimation

To begin with, some metrics for assessing early snow melt and pollinator diversity do produce significant effects. For example, estimating the effect of early snow melt using treatment as an IV on the maximum diversity DOY produces a significant negative result. This suggests that early snow melt as effected by black sand application advances maximum diversity DOY by about 13 days, depending on the quantifier of diversity (Simpson's or richness; see Table 2). Encouragingly, these relationships hold in order of magnitude and direction when estimating the ATE, suggesting that non-compliance is not a dramatic issue and that our IV is reliable.

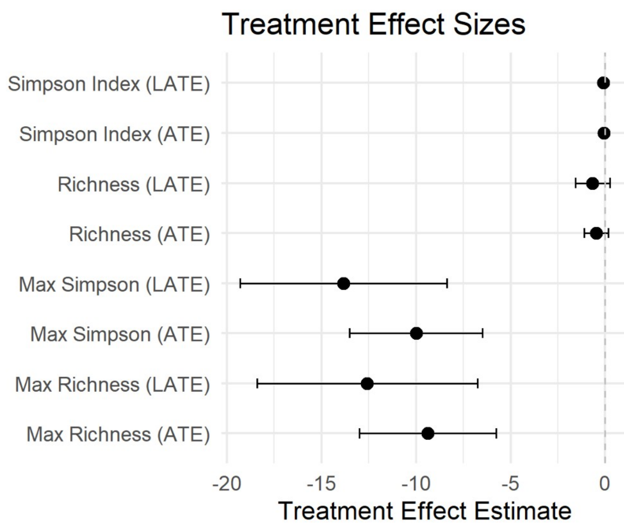

Figure 5. Effect sizes of various estimates of the LATE and ATE considered as total subplot diversity (Simpson's diversity index or species Richness) or Max diversity DOY. All estimates are significant except for the models using total Richness.

While these results are promising and match the findings of Rose-Person et. al., 2024b, they are not the whole story. For one, if treatment is reliably shifting snow melt, and snow melt is reliably shifting peak pollinator diversity, it would be expected for there to be a change when pollinator diversity is plotted over time in treatment versus control subplots. However, according to the Diversity Comparison ANOVA, this shift is not observed (Figure 4). This smoothed spline prediction takes into account the change in diversity over timeAdditionally, the sample size for estimations of the LATE are very small (n = 24) and the variance in snow melt DOY very low. Data would need to be collected over multiple years to determine if this trend is consistent.

Conclusion

While using correctly specified causal assumptions did produce models with consistent and significant estimands, the nature of these assumptions must still be reviewed. The Niwot Ridge LTER Black Sands Experiment does provide instructive information for studying warming in high alpine ecosystems, the results from this experiment may not be consistent enough to draw generalizable or transferable conclusions. For one, the assumption is made that the effect being estimated is that of early snow melt on pollinator diversity, not the effect of black sand on pollinator diversity. If black sand application is affecting pollinator diversity through other causal pathways that are not accounted for by the year fixed effect, this assumption may be violated by omitted variable bias. Another concern is the limited number of pollinator visitation data collected, spanning a single season. While the LATE was estimated by omitting non-compliant subplots, did these treatment subplots melt earlier than controls because of the artificial intervention, or some other inter-year variance? If the latter, the controls would not be serving as appropriate counterfactuals to treated subplots.

Further avenues for assessing this research question include different quantification methods for both early snow melt and pollinator diversity, as well as other drivers of pollinator network diversity. For example, snow melt could be quantified as the rate of snow depth change over time, and its effects on flower diversity and abundance could be investigated. While a consistent negative effect of early snow melt on various diversity metrics was seen over this analysis, there may be better methods for quantifying these metrics. The nature of the definitions of early snow melt and pollinator diversity could be catering to a desired estimate of the effect. As shown by the GAMs, day of peak diversity may not be reflective of the temporal dynamics of pollinator interaction events.

In conclusion, the results of this experiment provide some evidence that early snow melt decreases pollinator diversity in high alpine systems, in agreement with Rose-Person et. al., 2024b. However, this analysis shows that the experimental design may need to be improved before these results can be generalized or transferred to other systems.

Works Cited

Forrester, C., J. Morse, and Niwot Ridge LTER. 2024. Soil moisture and snowdepth measurements in the Black Sand experiment for East Knoll, Audubon, Lefty, Soddie and Trough, 2018 - 2024. ver 9. Environmental Data Initiative. https://doi.org/10.6073/pasta/d0f0eb82c4dfd90b21270710c82b16ce (Accessed 2025-04-29).

Rose-Person, A., M. Spasojevic, C. Forrester, W. Bowman, K. Suding, M. Oldfather, and N. Rafferty. 2024. Pollinator visitation, flower count, and seed set in Black Sand plots, 2020. ver 4. Environmental Data Initiative. https://doi.org/10.6073/pasta/ee4298a90e8947a0402123d2caf16f5c (Accessed 2025-04-29).

Rose-Person, A., Spasojevic, M. J., Forrester, C., Bowman, W. D., Suding, K. N., Oldfather, M. F., & Rafferty, N. E. (2024). Early snowmelt advances flowering phenology and disrupts the drivers of pollinator visitation in an alpine ecosystem. *Alpine Botany*, *134*(2), 141–150. https://doi.org/10.1007/s00035-024-00315-x

Data and Analysis Transparency

All data and R code used in this analysis and for the generation of figures can be found here.