The Pennsylvania State University
The Graduate School

A BAYESIAN SEGMENTED LINEAR REGRESSION MODEL TO OBSERVE BREAKFAST CANYON PATTERNS IN HONEYBEE HIVE WEIGHT DATA

A Thesis in
Statistics
by
Francesca Yaukey

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Master of Science

December 2023

The thesis of Francesca Yaukey was reviewed and approved by the following:

Ephraim Hanks
Associate Professor of Statistics
Thesis Co-Advisor

Stephen Berg
Assistant Professor of Statistics
Thesis Co-Advisor

Stephen Berg
Assistant Professor of Statistics
Committee Member

Bing Li
Professor of Statistics
Head of the Graduate Program

Abstract

Honeybees are essential to ecosystem health and crop vitality, but unfortunately are threatened by large scale changes in land use, increased pesticide use, and industrialized agricultural practices. Better understanding of honeybee colony health and honeybee behavior can lead to better management of bee colonies, enabling bee researchers to help combat colony collapse disorder (CCD). Utilizing hive weight data, we wish to explore the existence of Breakfast Canyons, an ecological phenomena that can give insight into the foraging force of a hive. We create a Bayesian segmented linear regression model which allows for multiple breakpoints to indicate where the slope changes in the daily weight trends. We apply our model to individual continuous hive weight data collected from honeybee colonies in Michigan collected throughout 2018, and explore the existence of the Breakfast Canyon, following a formally established definition of a Breakfast Canyon. Further, we establish a new definition of a Breakfast Canyon based on relative slope patterns, and explore the existence of a Breakfast Canyon with the same data. To further establish the existence of the Breakfast Canyon, we apply our methods for Breakfast Canyon definition utilizing both Meikle's absolute slope definition and our new relative slope definition to simulated Brownian Motion data, and compare the results with the real Honeybee data results.

Table of Contents

List of Figures

List of Tables

Acknowledgments

I extend my sincere gratitude to my co-advisors, Dr. Ephraim Hanks and Dr. Stephen Berg, for their invaluable guidance and unwavering support throughout my research journey. Their assistance was particularly crucial during periods of slow progress due to life's challenges. Without their support, feedback, and belief in my capabilities, navigating this academic undertaking would have been significantly more difficult. I am profoundly thankful to Dr. Christina Grozinger, Darcy Gray, Dr. Gabriela Quinlan, and the Grozinger Lab for graciously allowing me access to the field of Honeybee Ecology. Their explanations of intricate concepts and provision of essential data were instrumental in shaping this master's thesis. Funding for this work was partially supported by the National Science Foundation through NSF DMS-2015273. The findings and conclusions do not necessarily reflect the view of the funding agency.

A special acknowledgment is due to my husband, Noah, whose steadfast support and encouragement paved the way for me to pursue my academic aspirations. His unwavering commitment, love, and patience have been the bedrock of my strength, enabling me to successfully complete this academic endeavor. His own motivation, determination, and success in law school has acted as a driver for me to finish my master's degree. Equally deserving of recognition is my son, a perpetual source of inspiration and happiness, motivating me towards continual personal and professional development. My husband and son's unwavering support, warm smiles, and comforting hugs have been indispensable to my academic achievements.

Heartfelt thanks are extended to my parents for their enduring support, from encouraging my pursuit of an undergraduate degree in Mathematics to providing childcare assistance during the demanding phases of both undergraduate and graduate studies. Gratitude is also extended to my siblings, Donovan and Jessica, as well as my sibling-in-law, Jeremy, for their steadfast support, both in person and over the phone, during challenging times. I express my appreciation to Kaitlyn Fales for her indispensable assistance with coursework, insightful feedback, and unwavering moral support and friendship throughout this challenging academic pursuit. Finally, I extend my sincere thanks to Dr. Karen Anderson, Dr. Sharon Galvin, MaeLena Grazier, Sara Bowman, and the entire HEALS team at UHS for their support through challenging times. I am thankful for my extensive support system's unwavering belief in my abilities, which significantly contributed to maintaining my spirits and motivation throughout this rigorous academic journey.

Chapter 1 | Introduction and Background

Honeybees serve as essential pollinators, and are critical to maintaining ecosystem biodiversity, and sustaining crop health. In this masters' thesis, we develop statistical methods to aid apiarists in the assessment of Honeybee colony health, with an emphasis on the use of weight scales to explore the Breakfast Canyon phenomenon. These methods can offer valuable insights into the foraging force strength of bee colonies, a critical metric to assess Honeybee colony health and activity. The understanding of foraging force size is vital for understanding internal and external factors important to Honeybee Ecology, addressing challenges they face such as Colony Collapse Disorder, and ensuring the preservation of their role in sustaining ecosystems and the global food supply. Honeybees' roles include maintaining ecosystem biodiversity, sustaining crop health, and serving as essential pollinators.

Meikle and Holst (2018) used Honeybee hive weight data, where weight observations were collected every 15 minutes. Honeybee hive weight data can give insights into Honeybee activity and behavior without disrupting the hive. In our thesis, we utilize Honeybee hive weight data collected from hives located in Michigan, USA and implement Meikle's absolute change definition of the Breakfast Canyon, founded on an absolute change slope pattern framework in a Bayesian Segmented Linear Regression model. Further, we introduce a novel definition of the Breakfast Canyon based on relative change slope patterns, allowing for the the observation of Breakfast Canyons that previously would not have been observed under Meikle's framework. Our work also employs criteria beyond R-squared to determine where a 2 or 3 break Segmented Linear model is a better fit when fitting within the Segmented Linear Regression model. By establishing these statistical methods and Breakfast Canyon definitions, this thesis contributes to the advancement of understanding Honeybee colony health and activity, ultimately benefiting the preservation and well-being of Honeybee populations and the ecosystems they sustain.

In Chapter 1, we discuss the importance of Honeybees ecologically, highlighting their role in pollination, their biodiversity and ecosystem importance, as well as the threats that face Honeybee populations. We outline the ecological factors and technologies used in Honeybee hive health and behavior analysis. We detail the Breakfast Canyon phenomena, which is a metric used to analyze the foraging force of Honeybee hives, a key indicator of Honeybee health and activity. Lastly, we talk about previous work around the Breakfast Canyon led by Hambleton (1925), Meikle, and Holst (2018). Later, in Chapter 2, we outline the Bayesian Segmented Linear Regression Model framework and Breakfast Canyon Definitions. Lastly, in Chapter 3, we apply our methods to real Honeybee hive data and data simulated from Brownian Motion.

1.1 Importance of Honeybees Ecologically

1.1.1 Role in pollination

Pollination is important in maintaining the balance of ecosystems and is a foundation of crop production and vitality. Pollination, the process of pollen moving from the male anthers to the female stigmata between plants or between the same flower, can be contributed to by animal pollinators, including bees, and weather phenomena such as wind and air (S. A. Khalifa et al, 2021). According to S.A. Khalifa et. al, 87 of the leading global food crops depend on animal pollination, while 28 crops do not. Global crop production is directly dependent on honeybee pollination, as honeybees are responsible for the pollination of approximately 35% of global food crops (A. Klein et al., 2006). Stated by Southwick & Southwick in their 1992 paper *Estimating the Economic Value of Honeybees (Hymenoptera: Apidae) as Agricultural Pollinators in the United States*, without honeybees, yields of certain fruit, nuts and seed crops decrease more than 90%. Southwick & Southwick claim that Honeybees are the most economically valuable, convenient, and affordable pollinators, providing crop vitality worldwide. Often, they state, the implementation of honeybees near agricultural fields is the only affordable solution for farmers to ensure their crops are pollinated.

1.1.2 Biodiversity and Ecosystem Importance

Honeybees also play a key role in plant biodiversity conservation. Biodiversity includes the diversity within and between species as well as the diversity of ecosystems, which is essential for next generation stability (Akhila A. et al. 2022). Honeybees provide a range of services for the ecosystem, along with pollination, that contribute to the well-being of people, and assist in maintaining the planet's ecosystem and achieving global sustainable development and maintenance (Akhila A. et al. 2022). Honeybees contribute to the maintenance of ecological diversity by pollinating a diverse range of wild plants (Steffan-Dewenter and Tscharntke, 2000). The importance of bee pollination on crop vitality and biodiversity has been studied widely, and various areas of the effects of honeybee pollination as well as honeybee products continue to be researched. To demonstrate a few examples, Honeybee pollination contributes to enhanced nutritional value, longer shelf life and better quality of fruits and vegetables, production of biofuels, urban biodiversity efforts, and livelihood improvements in areas like Africa (Akhila A. et al. 2022). The presence of honeybee populations overall can enhance the stability and resilience of ecosystems, as well as contribute to the sustainable development of societies (R. F. Mortiz et al., 2007).

1.2 Threats to Honeybee Populations and the Importance of Analysis

In recent years, there has been a decline in honeybee populations and potential drivers of bee colony decline, and the associated impacts on the ecosystems are not fully understood (Akhila A. et al. 2022). In the U.S between April 2019 and April 2020, a reported 43% of bee colonies were lost, caused by several factors, either alone or in combination (Insolia et al., 2022). Honeybee colony loss rates are influenced by many factors, including parasite and pathogen levels when combating pests such as the *Varroa destructor* (Insolia et al., 2022). Diseases like Nosema and *Varroa destructor* mites can devastate honeybee colonies. The *Varroa destructor* mite especially has spread rapidly worldwide within a short time period, and without periodic treatment can cause colonies to collapse within a 2–3 year time period (Rosenkranz et al., 2010). Further, how the surrounding land is utilized (i.e., urbanization, and agricultural land use) can impact the foraging quality and the honeybees' exposure to pesticides (Insolia et al., 2022). Honeybees are vulnerable to pesticides, which can have detrimental effects on their health and behavior (Sanchez-Bayo

& Goka, 2014).

Biodiversity threats like climate change, agricultural expansion and invasive species further threaten honeybee distribution and abundance (Insolia et al., 2022). Colony failure often includes overwintering losses in cold climates and loss due to diseases during the winter. Losses due to these weather related stressors have increased with climate change (Insolia et al., 2022). The driver of honeybee colony loss can be puzzling, and previously healthy hives have been found abandoned, the hives absconded. The hives still contain stored food, dead brood, and a few dead bees, with no clear explanation for this rapid absconding from the hive (Myerscough et al., 2016). Generally, this phenomenon occurs in the late spring or the early summer, right as the hives emerge from their winter hibernation and reach peak numbers during the summer. This phenomenon, which is being widely researched, is known as colony collapse disorder (CCD) (Myerscough et al., 2016). CCD has led to the alarming decline in honeybee populations, posing a significant threat to agriculture and food security (van Engelsdorp et al., 2009).

Given the decline in honeybee populations along with the threats they face such as land use changes, urbanization, agriculture intensity, climate change, *Varroa destructor* and other diseases, pesticides, and colony collapse disorder, there is a critical need for quantitative approaches to analyze honeybee colony health, performance, and behavior, and thus protect biodiversity, ecosystem health, and crop vitality.

1.3 Analyzing Honeybee Hive Health and Behavior

Monitoring and data collection of ecological factors is imperative in the realm of beekeeping and apiary research. Factors such as hive weight, bee activity, levels of brood, honey, or pollen, internal hive temperature and weather conditions can each offer valuable insights into honeybee colony health and behavior. In this section, we will briefly explore these ecological factors, as well as provide an overview of the tools and technologies that apiarists and researchers use to collect this data.

1.3.1 Ecological Factors in Hive Health and Behavior Analysis

There are numerous ecological factors that apiarists and researchers would like to monitor for hive health analysis. Hive weight is an important factor in assessing colony health, as changes in the hive weight can indicate changes in food stores, honey production, brood production, and fluctuations in the population. Internal hive temperature is important to monitor, as temperature fluctuations in the hive can affect bee foraging behavior, development of brood, and overall hive survival. While difficult to monitor, bee activity, such as when bees leave and re-enter the hive can give researchers helpful insights into foraging activity, brood production, and communication between Honeybees. The levels of honey, brood, and pollen are also difficult to monitor, but can provide insight into hive nutrition, population growth, and overall hive health. Externally, weather data is important to monitor, as honeybee behavior and activity could depend on weather phenomena such as temperature, humidity, precipitation, solar radiation, and wind (e.g., Karbassioon et al., 2023).

1.3.2 Technologies Utilized in Hive Health and Behavior Analysis

In this section, we will discuss different technologies utilized in hive health and behavior analysis, starting with in-hive sensors. Implementation of electronic and digital sensor technologies in apiary research and general beekeeping has further improved understanding of Honeybee behavior and activity. Temperature sensors, as well as humidity sensors in bee hives can give apiarists valuable information about how the Honeybees are regulating their in-hive temperature and humidity to aid in health queen egg production, food preservation, and brood rearing (Zaman et. al, 2023). They highlight that research has found that using temperature sensors monitoring honeybee colony brood temperature can be an indicator of when a colony may swarm. Further, they discuss that certain in-hive sensors can monitor acoustics within the hive and evidence has been found that brood cycles may be related to vibrations within a cell. Other types of sensors that may be beneficial include gas sensors, thermal imaging technology, infrared sensors, vibration detectors, accelerometers, and light intensity sensors. Overall, in-hive sensors can provide valuable insights to apiarists and researchers, while avoiding disrupting the honeybees or the hive. However, these sensors do not give direct insight into when bees are leaving to forage and returning from foraging, which is a pivotal metric in assessing the overall health of the colony.

Radio-frequency identification (RFID) systems are another emerging technology used in hive health and behavior analysis. Streit et al. introduced radio-frequency identification tags (RFID) that monitor honeybee activity, particularly honeybee movement in and out of the hive (2003). As detailed in the paper, the small and light tags are attached to adult bees, and a detector records each time the bee with the attached tag crossed a threshold. While RFID tag systems are useful tools to examine bee foraging and foraging capacity, individual hive dynamics, and bee life expectancy, they are difficult to implement as implementation involves individually tagging the bees, which can distress them and has a potential to cause the bees harm (Nunes-Silva, 2018). However, these sensors give direct insight into when bees leave and return from foraging. While we are focusing on foraging in the scope of this masters' thesis, we will not be utilizing data collected from RFID technology because of these difficulties.

Lastly, we discuss the use of hive weight scales in hive health and behavior analysis. Newer approaches in apiculture include high resolution monitoring of beehive weight over time via the use of beehive scale weights. The collected hive weight data give apiarists and researchers insight into honeybee colony health, behavior, and performance (Schmickl & Crailsheim, 2004). Further, scales provide valuable data on honey storage, brood production, and overall colony growth (Seeley, 1982). The raw weight data is a function of several factors over time, including colony food collection and consumption, moisture gain due to humidity and rain, overnight evaporation of moisture, honeybee breeding and disappearance, colony robbing, absconding, and swarming (Meikle, Holst et al., 2018). While raw weight data is complex, analysis of such has shown to provide information on effects of weather on honeybee colonies, evidence towards causes of swarming and hive abandonment, differences in patterns among honeybee races, colony growth and consumption, overwintering, and the impact of pesticides on bee colonies. Observation of long-term colony weight changes allows for analysis of gains and losses of honey and pollen stores as well as bee population over time. Observation of day-to-day weight changes allow for analysis of departure and return of foragers, influx of resources, changes in humidity and moisture content of food and the hive structure, and abrupt changes to weight due to disturbance to the hive. The data we will be utilizing in this masters' thesis is hive weight data, which we will further explain in the data description section.

1.4 The Breakfast Canyon Phenomena

The Breakfast Canyon phenomenon is an ecological concept in Honeybee colony hives, which is well-known but is less understood. The Breakfast Canyon offers a valuable metric for evaluating the foraging performance and capacity of individual honeybee colonies. The Breakfast Canyon depth serves as an indicator of a colony's foraging force. A lower Breakfast Canyon depth corresponds to a smaller foraging force, while a higher Breakfast Canyon indicates a larger foraging force. Foraging force, which can give insight into the efficiency of a hive's food collection, can provide insights into the foraging capabilities, as well as the nutritional health of a Honeybee hive with minimal hive intervention, as it relies on non-intrusive hive weight data. Further understanding of the existence and significance of the Breakfast Canyon phenomenon is crucial to assist apiarists and ecological researchers, given the pivotal role of the foraging force in analyzing Honeybee colony health and long-term performance (Colin, 2022). The size of a colony's foraging force directly influences food collection and hive maintenance, making it a key determinant of colony vitality (Seeley, 1995).

In this masters' thesis chapter, we discuss foraging force and the importance of further research to understand the foraging force as a metric of Honeybee hive health. We discuss previous methods to analyze foraging force, including RFID technologies and the use of Honeybee hive weight data and the observation of the Breakfast Canyon phenomenon. Further, we discuss previous work by Hambleton, Meikle, and Holst on observing the Breakfast Canyon phenomenon in Honeybee hive weight data.

1.4.1 Analyzing the Foraging Force of Honeybee Hives

Honeybee foragers have an important role in the colony, as they must supply their colony with pollen and nectar. Without foragers, honeybees would have no chance of maintaining sustenance and health, resulting in colony failure (Klein, 2019). A key ecological concept in honeybees is the foraging force size, which indicates the size of the foraging population, as well as the availability to forage. Foraging force is a critical factor in honeybee colony health, as it directly influences food collection and hive maintenance (Seeley, 1995). The long-term performance of a colony is dependent on how capable the foraging force is, with colonies with high foraging performance resulting in thriving hives (Colin, 2022). Observing the size of a colony's foraging force is a useful tool in analyzing a colony's foraging performance, however there is a current need for more datasets and methods in observing foraging force (Colin, 2022).

Additional methods for analyzing honeybee foraging performance are crucial to aid in the understanding of colony collapse. Further, methods for analyzing honeybee performance could result in critical information needed for making impacts in environmental policy-making (Colin, 2022). Foraging force and the associated foraging performance have been researched through a few methods, including RFID analysis (Klein et. al., 2019). Klein et. al. used radio frequency identification (RFID) technology to observe the arrival and departure of individual bees through the colony entrance, as well as the length of the individual bee's foraging trips (2019). This provided valuable information on foraging activity and performance and found that a minority of very active bees (19% of foragers) performed 50% of the individual colony's total foraging trips (Klein et al., 2019). As established earlier in this masters' thesis, RFID technology is useful, but difficult to implement and disruptive to the hive.

1.4.2 Breakfast Canyon Description and Background

A less disruptive and more cost-effective method to analyze the foraging performance and capacity is proposed by Holst and Meikle (2018). They propose that we may observe the depth of an ecological occurrence deemed the Breakfast Canyon, which we will describe further in the methods section of this paper. Holst and Meikle (2018) suggest that the Breakfast Canyon (BC) depth can give valuable information on the foraging performance and foraging capacity of individual bee colonies, with low BC depth translating to a smaller foraging force, and high BC depth translating to a larger foraging force. The Breakfast Canyon (BC) depth measures the efficiency of food collection and can indicate the foraging force of a hive (Seeley, 1995). Meikle's approach in estimating the foraging force of individual colonies relies on honeybee hive weight data, which is less intrusive to bees than RFID, and is easier to implement. Understanding Breakfast Canyon depth could provide insights into the foraging capabilities and nutritional status of a honeybee colony, without minimal hive intervention (Seeley, 1995). While the idea of approximating foraging force from a Breakfast Canyon is appealing, the prevalence of this Breakfast Canyon phenomena is still poorly understood.

1.4.3 Previous Work and Extensions

The existing definitions of a Breakfast Canyon are vague, but primarily revolve around the existence of an early morning dip in the honeybee hive weight. Hambleton (1925) first identified a daily Beehive weight pattern, which was displayed as a series of line segments that had varying slopes. Hambleton identified three line segments, including a segment with a negative slope from 5am to 9am, followed by a positive slope from 9am to 7pm, then again followed by a negative slope from 7pm overnight to the next day at 5am. Hambleton identified the negative slope from 5am to 9am as the "morning loss" that occurs as the bees are leaving the hives to forage for the day, reaching the lowest point at the now named Breakfast Canyon.

Meikle and Holst (2018) expanded upon Hambleton's idea and found a pattern in their Beehive weight data in Southern Arizona, where they observed a distinct dip in weight in the mornings, which they named the Breakfast Canyon. They later realized that the low point in the canyon corresponded to the low point in Hambleton's approach. Their approach to describe the daily weight pattern identified was to use a segmented linear regression procedure, which estimates n linear regression lines joined at n-1 breakpoints. They found that an Segmented Linear Regression (SLR) with n=4 segments and 3 breakpoints to have the best fit, using R-squared as the criteria to pick the best number of segments to fit. They estimated the nocturnal loss rate as the negative slope of the longest segment ending at or before the rim of the Breakfast Canyon. From there, they also estimated canyon depth of the SLR models that were identified to have a breakfast canyon. They fit 3,727 SLRs, and of those 1,904 allowed for estimation of the canyon depth and nocturnal loss rate. Their primary findings were that the early morning dip in the weight (Breakfast Canyon) could be extracted from their data to then provide information on foraging effort and foraging force of honeybee colony performance.

We develop statistical methods to aid beekeepers in assessing honeybee colony health, with a particular emphasis on weight scales and analysis of the Breakfast Canyon phenomenon for insights into foraging strength. Building upon Holst and Meikle's segmented linear regression approach, we explore Meikle's concept of a breakfast canyon based on absolute change slope patterns, and introduce a new definition based on relative change slope patterns. Additionally, we aim to identify the optimal number of segments for the segmented linear regression model using criteria other than R-squared. The R-squared value for a 4 segment SLR fit to a day will always be higher than a 3 segment SLR fit to the same day. 4 segment SLR models have more parameters than 3 segment SLR models, and R-squared will be higher when you add more parameters. Thus, if we were to use R-squared to choose between a 4 segment SLR and a 3 segment SLR, the 4 segment SLR will be chosen due to the higher number of parameters. We establish a Bayesian segmented linear regression model which allows for multiple breakpoints to indicate where the slope changes in the daily weight trends, as well as allow for a more principled model selection technique by utilizing WAIC. We apply our model to individual continuous hive weight data collected from honeybee colonies in Michigan collected throughout 2018 and explore the proportion of days in which there is a breakfast canyon present using both Meikle's definition, as well as our new definition.

Chapter 2
A Bayesian Segmented Linear Regression Model Framework and Methods

In this Master's thesis chapter, we describe a new method for observing Breakfast Canyons in honeybee hive weight data. We detail the framework of both a 3 breakpoint and a 2 breakpoint Bayesian Segmented Linear Regression model, as well as a Bayesian approach for the estimation of parameters, and a model selection technique to determine which model is a better fit for each hive-day case. We outline the existing Breakfast Canyon definition of Meikle and Holst for both a 2 breakpoint and 3 breakpoint SLR model, which is based off of absolute slope patterns and breakpoint timing. We introduce a novel extended Breakfast Canyon definition for both a 2 breakpoint and 3 breakpoint SLR model, which instead is based off of relative slope change patterns and breakpoint timing. In Chapter 3, we apply this approach to 3,930 daily time series of hive weights from Honeybee colonies in Michigan.

2.1 Bayesian SLR Model with 3 Breakpoints

We first lay out the framework for the 3 breakpoint segmented linear regression model. The segmented linear model we establish has n=4 segments and n-1=3 breakpoints, following Holst and Meikle (2018). The breakpoints represent the changes in the linear trend behavior of the honeybee hive weight data from 12am-11am. We begin describing the model by letting $T > 0$ be an integer indicating the number of observed times in a given data set. In our case, because we have 44 weight measurements at 15-minute increments from 12am-11am, T=44. Our observed data, which is the weight of

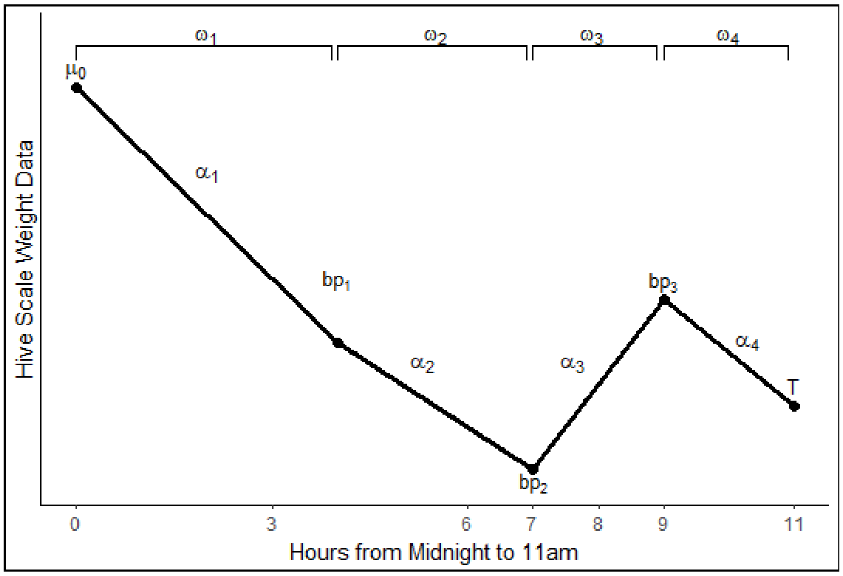

Figure 2.1: This figure gives a visualization of the parameters in the Bayesian 3 breakpoint Segmented Linear Regression model. $\alpha_1, \alpha_2, \alpha_3$ and α_4 are the slope parameters. μ_0 is the parameter for the y-intercept. $\omega_1, \omega_2, \omega_3$ and ω_4 are the parameters representing the spacing between the breakpoints. bp_1, bp_2 and bp_3 are the breakpoints parameters.

the hive after normalizing by subtracting the midnight weight, can be represented as $Y_t = \{Y_1, Y_2, ..., Y_{44}\}$. The SLR model with three breakpoints, denoted as bp_1, bp_2, and bp_3, is given by:

$$Y_t = mean_t + \epsilon_t$$

with:

$$mean_t = \begin{cases} \mu_0 + \alpha_1 * t & \text{if } t < bp_1 \\ \mu_0 + \alpha_1 * bp_1 + \alpha_2 * (t - bp_1) & \text{if } t \in [bp_1, bp_2) \\ \mu_0 + \alpha_1 * bp_1 + \alpha_2 * (bp_2 - bp_1) + \alpha_3 * (t - bp_2) & \text{if } t \in [bp_2, bp_3) \\ \mu_0 + \alpha_1 * bp_1 + \alpha_2 * (bp_2 - bp_1) + \alpha_3 * (bp_3 - bp_2) + \alpha_4 * (t - bp_3) & \text{if } t \in [bp_3, T) \end{cases} \tag{2.1}$$

Here, $\mu_0, \alpha_1, \alpha_2, \alpha_3$, and α_4 are the regression parameters and ϵ_t is the error parameter. We assume ϵ_t has a Normal distribution, with mean 0 and variance σ^2, where $\sigma^2 \sim exp(1)$. With three breakpoints, the parameters we need to estimate are μ_0, α_1, α_2, α_3, α_4, bp_1, bp_2 and bp_3. The estimation of the breakpoints has an additional step, which we will further detail in the parameter estimation section. For ease of notation we represent our regression parameters as $\underline{\beta}$ =(μ_0, α_1, α_2, α_3, α_4).

2.1.1 Estimation of the Parameters

We conduct inference on the parameters of the model using a Bayesian approach. The posterior distribution is $[\underline{\theta}|y] \propto [y|\underline{\theta}][\underline{\theta}]$ where $\underline{\theta} = (\underline{\beta}, bp_1, bp_2, bp_3, \epsilon_t)$. Here $[\underline{\theta}|y]$ is the posterior distribution, $[y|\underline{\theta}]$ is the likelihood function of the model (equation 2.1), and $[\underline{\theta}]$ is the prior distribution of $\underline{\theta}$. The Bayesian methodology used here intertwines the prior information with the likelihood function to obtain the joint posterior distribution of all parameters.

We assign each of the regression parameters $\underline{\beta} = (\mu_0, \alpha_1, \alpha_2, \alpha_3, \alpha_4)$ independent normal prior distributions with mean 0 and variance 100^2, i.e. $N(0, 100^2)$. The parameter ϵ_t is assumed to have a prior distribution that is Normal with mean 0, variance σ^2, where σ^2 comes from a prior distribution that is exponential with mean 1. To establish the prior distributions for the breakpoints, we define:

$$
\begin{aligned}
bp_1 &= \omega_1 * T \\
bp_2 &= (\omega_1 + \omega_2) * T \\
bp_3 &= (\omega_1 + \omega_2 + \omega_3) * T
\end{aligned}
$$

And specify a Dirichlet prior distribution for $\underline{\omega} = (\omega_1, \omega_2, \omega_3, \omega_4)'$ which is appropriate for the prior of our breakpoints as Dirichlet random variables sum to 1. We put a slightly informative prior on the distribution, namely:

$$
\begin{pmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \\ \omega_4 \end{pmatrix} \sim Dir(\underline{1} * v)
$$

So, the prior mean of $\underline{\omega}$ is $E(\underline{\omega}) = \frac{1}{\sum_{k=1}^{4} \omega_k}[\omega_1, \omega_2, \omega_3, \omega_4]' = \frac{1}{4}\underline{1}$. An increased v decreases the variance on the prior, making it more likely that the breakpoints are more evenly spaced. We choose $v = 3$, which results in a variance of $var(\omega_k) = \frac{\omega_k(\sum_{k=1}^{4} \omega_k - \omega_k)}{(\sum_{k=1}^{4} \omega_k)^2(\sum_{k=1}^{4} \omega_k + 1)} = \frac{3}{80}$.

2.1.2 Observing Breakfast Canyons in Data Following Meikle's Definition in a 3 Breakpoint SLR Model

Meikle and Holst (2018) described what combination of timing of the breakpoints corresponds to the Breakfast Canyon. The timing of the breakpoint that corresponds to the Breakfast Canyon depends on where geographically the hives are located, but generally occurs between 5am-10am. Patterns of the line segment slopes that define a Breakfast Canyon pattern are as follows. One of the line segments with a negative slope must be followed by one with a positive slope. The segment pattern can have a negative 2nd segment followed by a positive 3rd segment, or a negative 3rd segment followed by a positive 4th segment. The segment pattern cannot be a negative 1st segment followed by a positive 2nd segment since the 1st segment begins at midnight and it doesn't make sense ecologically for the Breakfast Canyon to commence at midnight, thus we ensure that the onset must begin after 3am. In the case where the 3rd segment has a negative slope, the 2nd segment has an even steeper and negative slope, and the 4th segment has a positive slope, the 2nd segment is indicated as the true onset of the Breakfast Canyon. In this instance, segments 2, 3, and 4 form the canyon, with the bottom between 3rd and 4th segments. The absolute change slope patterns that are possible using Meikle's 3 breakpoint Breakfast Canyon definition are summarized in table 2.1.

In a 3 breakpoint segmented linear regression model with 4 segments, 16 absolute change slope patterns are possible, shown in Figure 2.2. Of the 16 absolute change slope patterns, 8 have a Breakfast Canyon following Meikle's 3 breakpoint SLR model breakfast canyon definition, highlighted in red. 4 of the patterns with an observed Breakfast Canyon reach the deepest point of the canyon at the 2nd breakpoint, and the other 4 patterns reach the deepest point of the canyon at the 3rd breakpoint.

2.1.2.1 Estimating Breakfast Canyon Depth

Once we have estimated the slope parameters from the 3 breakpoint SLR model, notated as $\hat{\alpha_1}, \hat{\alpha_2}$, $\hat{\alpha_3}$, and $\hat{}$ and the posterior means of the breakpoints, notated as $\hat{bp_1}$, $\hat{bp_2}$, and $\hat{bp_3}$,we are able to calculate the Breakfast Canyon Depth for the days in which a Breakfast Canyon was observed. For the days that a Breakfast Canyon was not observed in the data, the Breakfast Canyon Depth was considered missing.

First we describe the calculation of the estimated Breakfast Canyon if the bottom of

Pattern	Segment 1	Segment 2	Segment 3	Segment 4
Pattern 1	$\hat{\alpha_1} > 0$	$\hat{\alpha_2} > 0$	$\hat{\alpha_3} < 0$	$\hat{\alpha_4} > 0$
Pattern 2	$\hat{\alpha_1} > 0$	$\hat{\alpha_2} < 0$	$\hat{\alpha_3} > 0$	$\hat{\alpha_4} > 0$
Pattern 3	$\hat{\alpha_1} > 0$	$\hat{\alpha_2} < 0$	$\hat{\alpha_3} > 0$	$\hat{\alpha_4} < 0$
Pattern 4	$\hat{\alpha_1} > 0$	$\hat{\alpha_2} < 0$	$\hat{\alpha_3} < 0$ and $\hat{\alpha_3} \leq \hat{\alpha_2}$	$\hat{\alpha_4} > 0$
Pattern 5	$\hat{\alpha_1} < 0$	$\hat{\alpha_2} < 0$	$\hat{\alpha_3} < 0$ and $\hat{\alpha_3} \leq \hat{\alpha_2}$	$\hat{\alpha_4} > 0$
Pattern 6	$\hat{\alpha_1} < 0$	$\hat{\alpha_2} < 0$	$\hat{\alpha_3} > 0$	$\hat{\alpha_4} > 0$
Pattern 7	$\hat{\alpha_1} < 0$	$\hat{\alpha_2} < 0$	$\hat{\alpha_3} > 0$	$\hat{\alpha_4} < 0$
Pattern 8	$\hat{\alpha_1} < 0$	$\hat{\alpha_2} > 0$	$\hat{\alpha_3} < 0$	$\hat{\alpha_4} > 0$
Pattern 9	$\hat{\alpha_1} > 0$	$\hat{\alpha_2} > 0$	$\hat{\alpha_3} < 0$	$\hat{\alpha_4} < 0$
Pattern 10	$\hat{\alpha_1} > 0$	$\hat{\alpha_2} > 0$	$\hat{\alpha_3} > 0$	$\hat{\alpha_4} < 0$
Pattern 11	$\hat{\alpha_1} > 0$	$\hat{\alpha_2} < 0$	$\hat{\alpha_3} < 0$	$\hat{\alpha_4} < 0$
Pattern 12	$\hat{\alpha_1} < 0$	$\hat{\alpha_2} > 0$	$\hat{\alpha_3} > 0$	$\hat{\alpha_4} > 0$
Pattern 13	$\hat{\alpha_1} > 0$	$\hat{\alpha_2} > 0$	$\hat{\alpha_3} > 0$	$\hat{\alpha_4} > 0$
Pattern 14	$\hat{\alpha_1} < 0$	$\hat{\alpha_2} > 0$	$\hat{\alpha_3} > 0$	$\hat{\alpha_4} < 0$
Pattern 15	$\hat{\alpha_1} < 0$	$\hat{\alpha_2} > 0$	$\hat{\alpha_3} < 0$	$\hat{\alpha_4} < 0$
Pattern 16	$\hat{\alpha_1} < 0$	$\hat{\alpha_2} < 0$	$\hat{\alpha_3} < 0$	$\hat{\alpha_4} < 0$
Pattern 17	$\hat{\alpha_1} < 0$	$\hat{\alpha_2} < 0$	$\hat{\alpha_3} < 0$ and $\hat{\alpha_3} > \hat{\alpha_2}$	$\hat{\alpha_4} > 0$
Pattern 18	$\hat{\alpha_1} > 0$	$\hat{\alpha_2} < 0$	$\hat{\alpha_3} < 0$ and $\hat{\alpha_3} > \hat{\alpha_3}$	$\hat{\alpha_4} > 0$

Table 2.1: Table the possible absolute change slope patterns possible from Meikle and Holst's definition of a Breakfast Canyon.

the Breakfast Canyon is at the second breakpoint. In this case, the onset of the Breakfast Canyon occurred at the first breakpoint, so we calculate the Breakfast Canyon Depth by subtracting the estimated hive weight at the second breakpoint, $\hat{bp_2}$, from the estimated hive weight at the first breakpoint, $\hat{bp_1}$. The estimated breakfast canyon depth in this case is notated as: $\hat{D} = mean_{\hat{bp_1}} - mean_{\hat{bp_2}}$

Next we describe the calculation of the estimated Breakfast Canyon if the bottom of the Breakfast Canyon is at the third breakpoint. In this case, the onset of the Breakfast Canyon could occur either at the first breakpoint, or the second breakpoint, depending on the absolute change slope patterns. In the instances where segments 2, 3, and 4 form the Breakfast Canyon, the onset of the breakfast canyon begins at the first breakpoint. Thus, the estimated Breakfast Canyon depth is calculated by subtracting the estimated hive weight at the third breakpoint, $\hat{bp_3}$, from the estimated hive weight at the first breakpoint, $\hat{bp_1}$. The estimated breakfast canyon depth in this case is notated as: $\hat{D} = mean_{\hat{bp_1}} - mean_{\hat{bp_3}}$

For the instances where the onset of the canyon begins at the second breakpoint, the

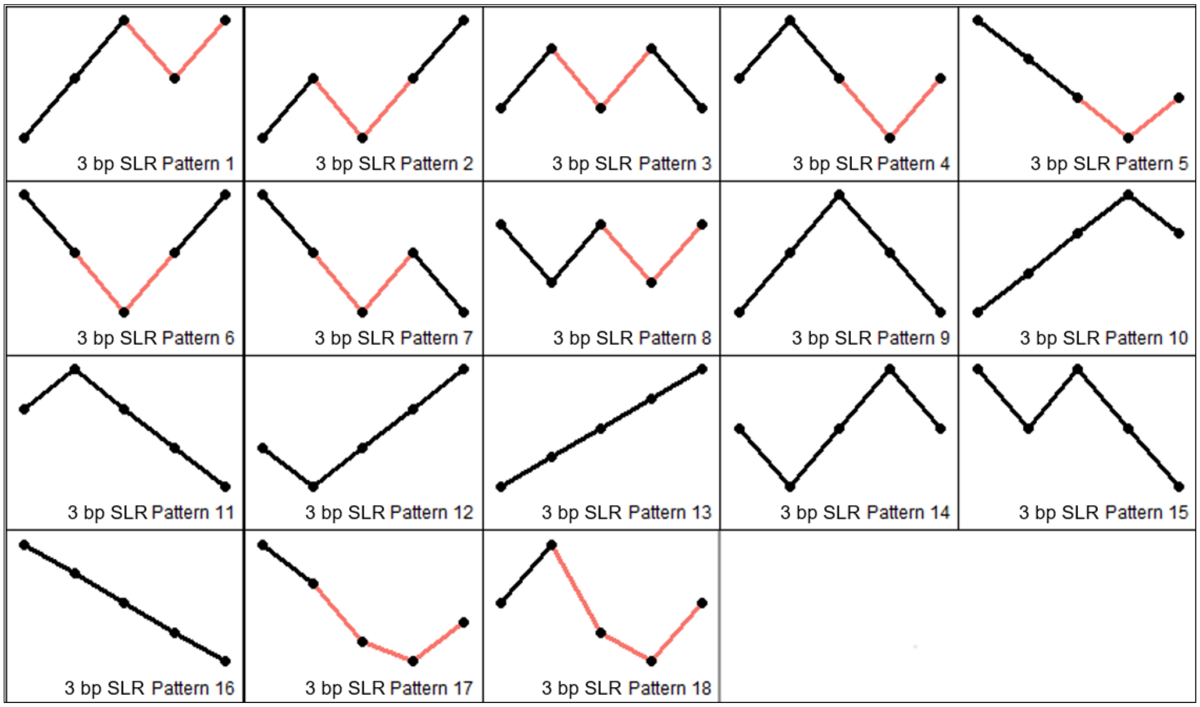

Figure 2.2: Pictured are the 18 absolute change slope patterns that are possible in a 3 breakpoint Segmented Linear Regression model. Patterns with segments highlighted red are patterns that have an observed Breakfast Canyon, following Meikle and Holst's definition of a Breakfast Canyon. The location of the breakfast canyon is highlighted in red. Patterns 17 and 18 are patterns in which the Breakfast Canyon consists of the 2nd, 3rd and 4th segments.

estimated Breakfast Canyon depth is calculated by subtracting the estimated hive weight at the third breakpoint, $\hat{bp}_3$, from the estimated hive weight at the second breakpoint, $\hat{bp}_2$. The estimated breakfast canyon depth in this case is notated as: $\hat{D} = mean_{\hat{bp_2}} - mean_{\hat{bp_3}}$

2.1.3 A Breakfast Canyon Definition Based Off of a Relative Slope Framework in a 3 Breakpoint SLR Model

Meikle's Breakfast Canyon Definition relies primarily on absolute slopes and is sub-setted within our new Breakfast Canyon definition based on relative slopes. Absolute slopes consist of only positive and negative slopes for the 3 or 4 segments of our Segmented Linear Regression model. Within this absolute slope framework, there is no comparison between the slope before or after the slope in question. Our innovative contribution lies in the introduction of relative slopes, which extends Meikle's definition to allow for the comparison between the strength of slopes of segments. The extension to the

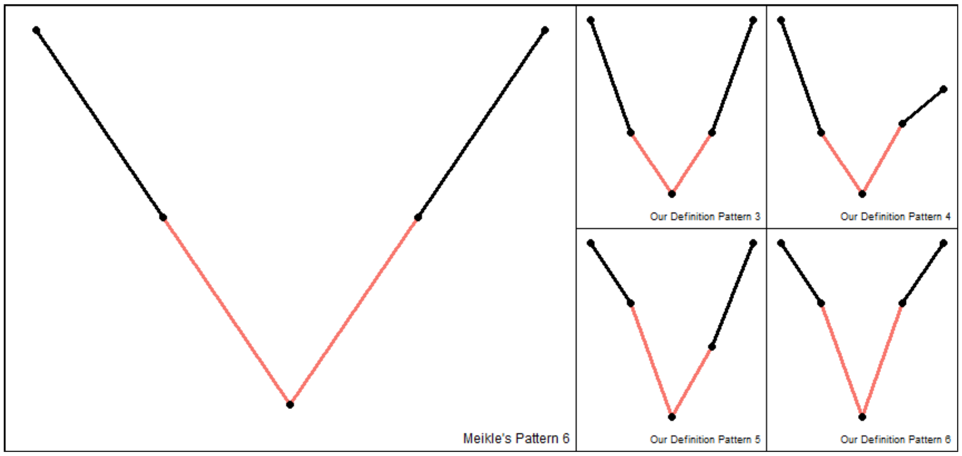

Figure 2.3: In this figure, we provide a visual of the breakdown of patterns that exist when switching from an absolute slope framework to a relative slope framework. Pattern 6 under Meikle's absolute change framework breaks down into four patterns (3,4,5 and 6) under the relative change framework.

relative slope framework adds additional patterns beyond Meikle's 3 breakpoint SLR model patterns.

For instance, if a pattern Meikle identifies has just two negative or two positive slopes in a row, two patterns are created in a relative slope framework. Further, if a pattern Meikle identifies has two consecutive positive slopes and two consecutive negative slopes, four patterns are created in a relative slope framework. A visualization of the breakdown of slope patterns is provided in Figure 2.3. If there are three negative slopes or three positive slopes in a row, 4 patterns are created in a relative slope framework. If there are four positive or four negative slopes in a row, 8 patterns are created in a relative slope framework. The 8 patterns Meikle identified as having an observed Breakfast Canyon extend with a relative slope framework to 18 patterns with a Breakfast Canyon. Further Breakfast Canyon patterns will be observed beyond these 18 patterns, which we move on to detail.

One of the advantages of our relative slope definition is the ability to identify Breakfast Canyons in scenarios where Meikle's definition may not result in a Breakfast Canyon Pattern, allowing us to observe potential Breakfast Canyons that otherwise would not be observed. For example, in Meikle's pattern 16, shown in Figure 2.4, where there are four negative slopes, notated as $(\hat{\alpha_1} < 0, \hat{\alpha_2} < 0, \hat{\alpha_3} < 0, \hat{\alpha_4} < 0)$, 8 patterns are created under a relative slope framework. There are many factors that could influence the overall

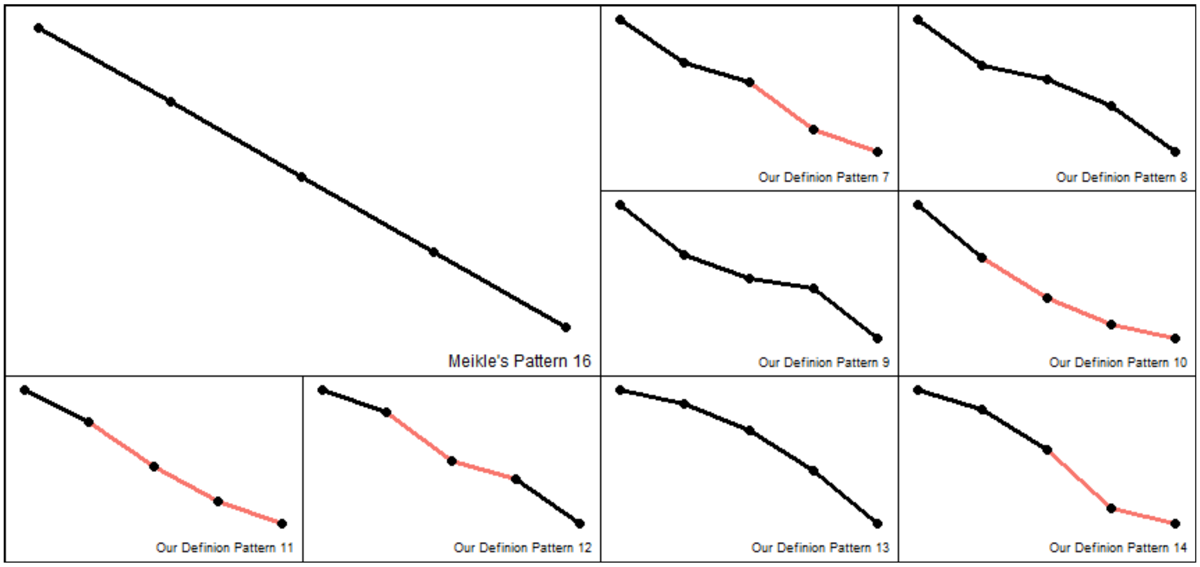

Figure 2.4: We provide a visualization of the additional patterns created from extending the Breakfast Canyon definition from Meikle's absolute slope framework to our relative slope framework, extending Pattern 16 into 8 different patterns.

weight trend of the honeybee hive weight. Imagine a day that is particularly dry and sunny, lacking humidity, that follows a day that had rain or significant humidity. As the moisture in the hive structure dries up, naturally the hive weight would decrease. Due to the already decreasing trend in the hive weight data, when the bees leave the hive to forage, the slope would become more negative. But, when the bees return from foraging, the change in weight may not be drastic enough to cause a positive slope, making the fourth slope negative. Under Meikle's definition, with four negative slopes, no Breakfast Canyon would be observed. However, under our relative slope change definition, when a negative first segment slope, followed by a negative second segment slope, followed by a third segment slope more negative than the second segment slope, then followed by a fourth segment slope less negative than the third, a Breakfast Canyon will be observed, such as in Pattern 7 and Pattern 14 in Figure 2.4.

Of the 8 patterns that Meikle identified as not having an observed Breakfast Canyon, 36 patterns were created under our relative slope definition. Of these 36 patterns, 17 had an observed Breakfast Canyon. 54 total patterns are created under a relative slope framework for a 3 break point SLR model, defined in tables 2.2 and .1.

Pattern	Segment 1	Segment 2	Segment 3	Segment 4
Pattern 1	$\hat{\alpha_1} < 0$	$\hat{\alpha_2} < 0$ and $\hat{\alpha_2} \geq \hat{\alpha_1}$	$\hat{\alpha_3} > 0$	$\hat{\alpha_4} < 0$
Pattern 2	$\hat{\alpha_1} < 0$	$\hat{\alpha_2} < 0$ and $\hat{\alpha_2} < \hat{\alpha_1}$	$\hat{\alpha_3} > 0$	$\hat{\alpha_4} < 0$
Pattern 3	$\hat{\alpha_1} < 0$	$\hat{\alpha_2} < 0$ and $\hat{\alpha_2} \geq \hat{\alpha_1}$	$\hat{\alpha_3} > 0$	$\hat{\alpha_4} > 0$ and $\hat{\alpha_4} > \hat{\alpha_3}$
Pattern 4	$\hat{\alpha_1} < 0$	$\hat{\alpha_2} < 0$ and $\hat{\alpha_2} \geq \hat{\alpha_1}$	$\hat{\alpha_3} > 0$	$\hat{\alpha_4} > 0$ and $\hat{\alpha_4} \leq \hat{\alpha_3}$
Pattern 5	$\hat{\alpha_1} < 0$	$\hat{\alpha_2} < 0$ and $\hat{\alpha_2} < \hat{\alpha_1}$	$\hat{\alpha_3} > 0$	$\hat{\alpha_4} > 0$ and $\hat{\alpha_4} > \hat{\alpha_3}$
Pattern 6	$\hat{\alpha_1} < 0$	$\hat{\alpha_2} < 0$ and $\hat{\alpha_2} < \hat{\alpha_1}$	$\hat{\alpha_3} > 0$	$\hat{\alpha_4} > 0$ and $\hat{\alpha_4} \leq \hat{\alpha_3}$
Pattern 7	$\hat{\alpha_1} < 0$	$\hat{\alpha_2} < 0$ and $\hat{\alpha_2} \geq \hat{\alpha_1}$	$\hat{\alpha_3} < 0$ and $\hat{\alpha_3} < \hat{\alpha_2}$	$\hat{\alpha_4} < 0$ and $\hat{\alpha_4} > \hat{\alpha_3}$
Pattern 8	$\hat{\alpha_1} < 0$	$\hat{\alpha_2} < 0$ and $\hat{\alpha_2} \geq \hat{\alpha_1}$	$\hat{\alpha_3} < 0$ and $\hat{\alpha_3} < \hat{\alpha_2}$	$\hat{\alpha_4} < 0$ and $\hat{\alpha_4} \leq \hat{\alpha_3}$
Pattern 9	$\hat{\alpha_1} < 0$	$\hat{\alpha_2} < 0$ and $\hat{\alpha_2} \geq \hat{\alpha_1}$	$\hat{\alpha_3} < 0$ and $\alpha \geq \hat{\alpha_2}$	$\hat{\alpha_4} < 0$ and $\hat{\alpha_4} \leq \hat{\alpha_3}$
Pattern 10	$\hat{\alpha_1} < 0$	$\hat{\alpha_2} < 0$ and $\hat{\alpha_2} \geq \hat{\alpha_1}$	$\hat{\alpha_3} < 0$ and $\hat{\alpha_3} \geq \hat{\alpha_2}$	$\hat{\alpha_4} < 0$ and $\hat{\alpha_4} > \hat{\alpha_3}$
Pattern 11	$\hat{\alpha_1} < 0$	$\hat{\alpha_2} < 0$ and $\hat{\alpha_2} < \hat{\alpha_1}$	$\hat{\alpha_3} < 0$ and $\hat{\alpha_3} \geq \hat{\alpha_2}$	$\hat{\alpha_4} < 0$ and $\hat{\alpha_4} > \hat{\alpha_3}$
Pattern 12	$\hat{\alpha_1} < 0$	$\hat{\alpha_2} < 0$ and $\hat{\alpha_2} < \hat{\alpha_1}$	$\hat{\alpha_3} < 0$ and $\hat{\alpha_3} > \hat{\alpha_2}$	$\hat{\alpha_4} < 0$ and $\hat{\alpha_4} \leq \hat{\alpha_3}$
Pattern 13	$\hat{\alpha_1} < 0$	$\hat{\alpha_2} < 0$ and $\hat{\alpha_2} < \hat{\alpha_1}$	$\hat{\alpha_3} < 0$ and $\hat{\alpha_3} \leq \hat{\alpha_2}$	$\hat{\alpha_4} < 0$ and $\hat{\alpha_4} \leq \hat{\alpha_3}$
Pattern 14	$\hat{\alpha_1} < 0$	$\hat{\alpha_2} < 0$ and $\hat{\alpha_2} < \hat{\alpha_1}$	$\hat{\alpha_3} < 0$ and $\hat{\alpha_3} < \hat{\alpha_2}$	$\hat{\alpha_4} < 0$ and $\hat{\alpha_4} > \hat{\alpha_3}$
Pattern 15	$\hat{\alpha_1} < 0$	$\hat{\alpha_2} > 0$	$\hat{\alpha_3} < 0$	$\hat{\alpha_4} < 0$ and $\hat{\alpha_4} > \hat{\alpha_3}$
Pattern 16	$\hat{\alpha_1} < 0$	$\hat{\alpha_2} > 0$	$\hat{\alpha_3} < 0$	$\hat{\alpha_4} < 0$ and $\hat{\alpha_4} \leq \hat{\alpha_3}$
Pattern 17	$\hat{\alpha_1} < 0$	$\hat{\alpha_2} > 0$	$\hat{\alpha_3} < 0$	$\hat{\alpha_4} > 0$
Pattern 18	$\hat{\alpha_1} < 0$	$\hat{\alpha_2} < 0$ and $\hat{\alpha_2} < \hat{\alpha_1}$	$\hat{\alpha_3} < 0$ and $\hat{\alpha_3} \geq \hat{\alpha_2}$	$\hat{\alpha_4} > 0$
Pattern 19	$\hat{\alpha_1} < 0$	$\hat{\alpha_2} < 0$ and $\hat{\alpha_2} < \hat{\alpha_1}$	$\hat{\alpha_3} < 0$ and $\hat{\alpha_3} < \hat{\alpha_2}$	$\hat{\alpha_4} > 0$
Pattern 20	$\hat{\alpha_1} < 0$	$\hat{\alpha_2} < 0$ and $\hat{\alpha_2} \geq \hat{\alpha_1}$	$\hat{\alpha_3} < 0$ and $\hat{\alpha_3} < \hat{\alpha_2}$	$\hat{\alpha_4} > 0$
Pattern 21	$\hat{\alpha_1} < 0$	$\hat{\alpha_2} < 0$ and $\hat{\alpha_2} \geq \hat{\alpha_1}$	$\hat{\alpha_3} < 0$ and $\hat{\alpha_3} \geq \hat{\alpha_2}$	$\hat{\alpha_4} > 0$
Pattern 22	$\hat{\alpha_1} < 0$	$\hat{\alpha_2} > 0$	$\hat{\alpha_3} > 0$ and $\hat{\alpha_3} > \hat{\alpha_2}$	$\hat{\alpha_4} < 0$
Pattern 23	$\hat{\alpha_1} < 0$	$\hat{\alpha_2} > 0$	$\hat{\alpha_3} > 0$ and $\hat{\alpha_3} \leq \hat{\alpha_2}$	$\hat{\alpha_4} < 0$
Pattern 24	$\hat{\alpha_1} < 0$	$\hat{\alpha_2} > 0$	$\hat{\alpha_3} > 0$ and $\hat{\alpha_3} < \hat{\alpha_2}$	$\hat{\alpha_4} > 0$ and $\hat{\alpha_4} < \hat{\alpha_3}$
Pattern 25	$\hat{\alpha_1} < 0$	$\hat{\alpha_2} > 0$	$\hat{\alpha_3} > 0$ and $\hat{\alpha_3} < \hat{\alpha_2}$	$\hat{\alpha_4} > 0$ and $\hat{\alpha_4} \geq \hat{\alpha_3}$
Pattern 26	$\hat{\alpha_1} < 0$	$\hat{\alpha_2} > 0$	$\hat{\alpha_3} > 0$ and $\hat{\alpha_3} \geq \hat{\alpha_2}$	$\hat{\alpha_4} > 0$ and $\hat{\alpha_4} < \hat{\alpha_3}$
Pattern 27	$\hat{\alpha_1} < 0$	$\hat{\alpha_2} > 0$	$\hat{\alpha_3} > 0$ and $\hat{\alpha_3} \geq \hat{\alpha_2}$	$\hat{\alpha_4} > 0$ and $\hat{\alpha_4} \geq \hat{\alpha_3}$

Table 2.2: Table the possible relative change slope patterns possible from our new definition of a Breakfast Canyon, with the first segment starting with a negative slope. The remaining 27 patterns that start with a positive slope can be found in Appendix A.

2.1.3.1 Estimating Breakfast Canyon Depth

The process to estimate the Breakfast Canyon Depth for a 3 breakpoint SLR model under our relative change slope pattern Breakfast Canyon definition is similar to the process defined utilizing Meikle and Holst's definition in section 2.1. For most patterns, the process of calculating the Breakfast Canyon depth will be the same. The patterns that will have to be approached differently are patterns in which there are four positive slopes. We will first detail the process of finding the Breakfast Canyon depth for the

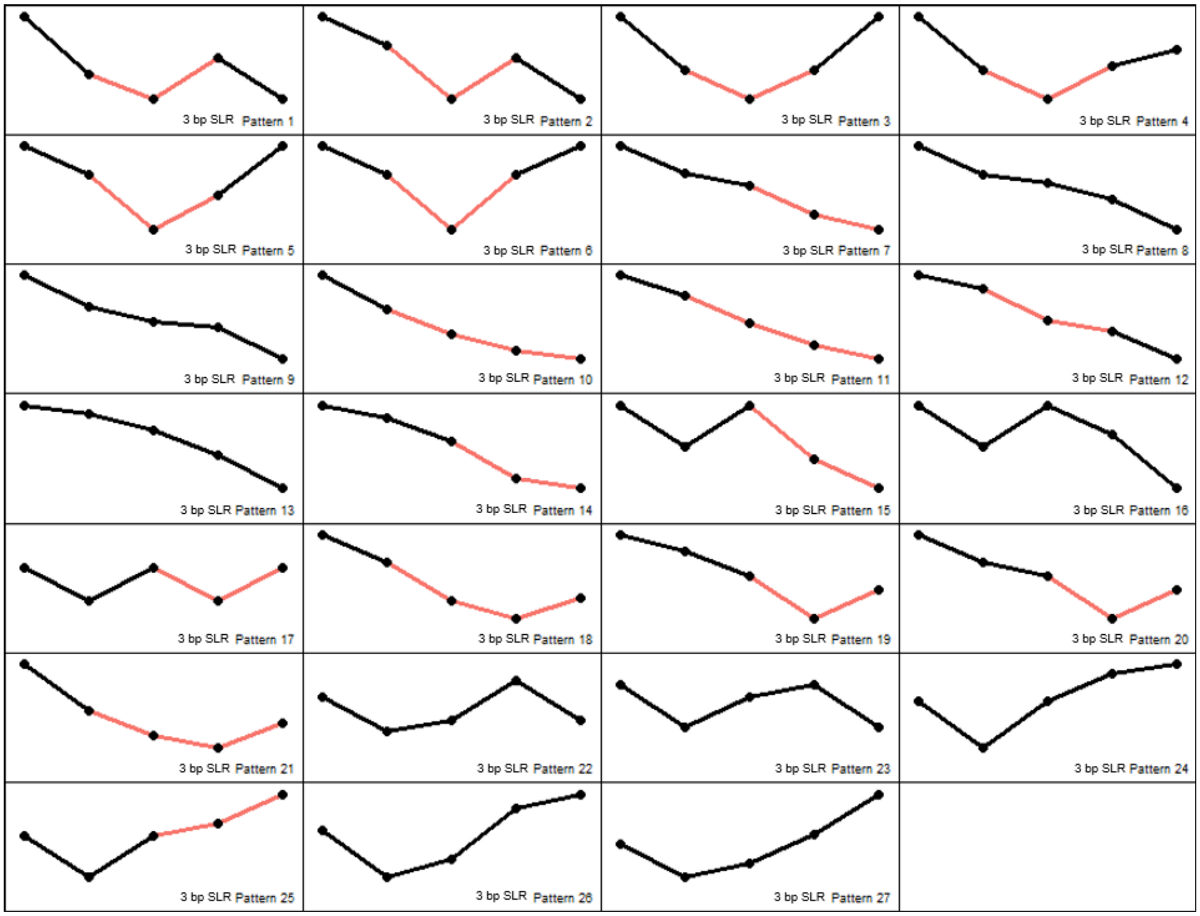

Figure 2.5: This figure gives a visualization that corresponds with table 2.2 of the first 27 of 54 possible relative change slope patterns using our 3 breakpoint Segmented Linear Regression model Breakfast Canyon definition .The remaining 27 patterns can be found in Appendix A.

patterns that will follow the same approach as Meikle's definition. Again, we have the posterior means of the slope parameters from the 3 breakpoint SLR model, notated as $\hat{\alpha_1}, \hat{\alpha_2}$, $\hat{\alpha_3}$, and $\hat{\alpha_4}$ and the posterior means of the breakpoints, notated as $\hat{bp_1}$, $\hat{bp_2}$, and $\hat{bp_3}$. Using the slope and breakpoint posterior means, we calculate the Breakfast Canyon Depth for the days in which a Breakfast Canyon was observed. If a Breakfast Canyon was not observed in the data, the Breakfast Canyon Depth was considered missing.

If the bottom of the Breakfast Canyon is at the second breakpoint, we calculate the Breakfast Canyon Depth by subtracting the estimated hive weight at the second breakpoint, $\hat{bp_2}$, from the estimated hive weight at the first breakpoint , $\hat{bp_1}$. The estimated breakfast canyon depth in this case is notated as: $\hat{D} = mean_{\hat{bp_1}} - mean_{\hat{bp_2}}$

If the bottom of the Breakfast Canyon is at the third breakpoint, again the onset of the Breakfast Canyon could occur either at the first breakpoint, or the second breakpoint, depending on the absolute change slope patterns. For Breakfast Canyon SLR models where segments 2, 3, and 4 form the Breakfast Canyon, the estimated Breakfast Canyon depth is calculated by subtracting the estimated hive weight at the third breakpoint, $\hat{bp}_3$, from the estimated hive weight at the first breakpoint, $\hat{bp}_1$. The estimated breakfast canyon depth in this case is notated as: $\hat{D} = mean_{\hat{bp_1}} - mean_{\hat{bp_3}}$

For 3 breakpoint SLR models where the onset of the canyon begins at the second breakpoint, the estimated Breakfast Canyon depth is calculated by subtracting the estimated hive weight at the third breakpoint, $\hat{bp}_3$, from the estimated hive weight at the second breakpoint, $\hat{bp}_2$. The estimated breakfast canyon depth in this case is notated as: $\hat{D} = mean_{\hat{bp_2}} - mean_{\hat{bp_3}}$

We now detail the process of calculating the Breakfast Canyon depth for the patterns in which there are four positive slopes. To calculate the Breakfast Canyon depth for these situations, the approach will be slightly different depending on which breakpoint the Breakfast Canyon reaches its' depth. First we will detail the process of calculating the Breakfast Canyon depth when the deepest point is at the third breakpoint. We find the distance between the weight at the time at which the Breakfast Canyon reaches its' deepest point if the slope of the third segment would have been the same as the second segment, and the weight at the time the Breakfast Canyon reaches its' depth following the actual point estimate of the weight at the deepest point of the Breakfast Canyon, $mean_{\hat{bp_3}}$. To find the expected weight at the time the deepest point of the Breakfast Canyon if the third segment slope was equivalent to the second segment slope, we calculate $\hat{Ybp}_3 = \hat{\alpha}_2 * t_{\hat{bp_3}} + c$ where $c = mean_{\hat{bp_2}} - \hat{\alpha}_2 * t_{\hat{bp_2}}$. The Breakfast Canyon depth in this case is notated as: $\hat{D} = \hat{Ybp}_3 - mean_{\hat{bp_3}}$.

Next we detail the process of calculating the Breakfast Canyon depth when the deepest point is at the second breakpoint. We find the distance between the weight at the time at which the Breakfast Canyon reaches its' deepest point if the slope of the second segment would have been the same as the first segment, and the weight at the time the Breakfast Canyon reaches its' depth following the actual point estimate of the weight at the deepest point of the Breakfast Canyon, $mean_{\hat{bp_2}}$. To find the expected weight at the time the deepest point of the Breakfast Canyon if the second segment slope was equivalent to the first segment slope, we calculate $Y\hat{bp}_2 = \hat{\alpha}_1 * t_{\hat{bp_2}} + c$ where $c = mean_{\hat{bp_1}} - \hat{\alpha}_1 * t_{\hat{bp_1}}$. The Breakfast Canyon depth in this case is notated as: $\hat{D} = Y\hat{bp}_2 - mean_{\hat{bp_2}}$.

2.2 Bayesian SLR Model with 2 Breakpoints

The 2 breakpoint segmented linear model we establish has n=3 segments and n-1=2 breakpoints. We let $T > 0$ be an integer indicating the number of observed times in a given data set with T=44, like the 3 breakpoint model. Our observed data, which is the weight of the hive after subtracting the midnight weight, can again be represented as $Y_t = \{Y_1, Y_2, ..., Y_{44}\}$.
The SLR model with two breakpoints, denoted as bp_1 and bp_2, is given by:

$$Y_t = mean_t + \epsilon_t$$

with:

$$mean_t = \begin{cases} \mu_0 + \alpha_1 * t & \text{if } t < bp_1 \\ \mu_0 + \alpha_1 * bp_1 + \alpha_2 * (t - bp_1) & \text{if } t \in [bp_1, bp_2) \\ \mu_0 + \alpha_1 * bp_1 + \alpha_2 * (bp_2 - bp_1) + \alpha_3 * (t - bp_2) & \text{if } t \in [bp_2, T) \end{cases} \tag{2.2}$$

Here, $\mu_0, \alpha_1, \alpha_2,$and α_3 are the regression parameters and ϵ_t is the error parameter. We assume ϵ_t has a Normal distribution, with mean 0 and variance σ^2, where $\sigma^2 \sim exp(1)$. With three breakpoints, the parameters we need to estimate are μ_0, α_1, α_2, α_3, bp_1, and bp_2. The estimation of the breakpoints has an additional step, which we will further detail in the parameter estimation section. For ease of notation we represent our regression parameters as $\underline{\beta}$ =(μ_0, α_1, α_2, α_3).

2.2.1 Estimation of the Parameters

We conduct inference on the parameters of the model using a Bayesian approach. The posterior distribution is $[\underline{\theta}|y] \propto [y|\underline{\theta}][\underline{\theta}]$ where $\underline{\theta} = (\underline{\beta}, bp_1, bp_2, \epsilon_t)$. Here $[\underline{\theta}|y]$ is the posterior distribution, $[y|\underline{\theta}]$ is the likelihood function of the model, equation 2.2, and $[\underline{\theta}]$ is the prior distribution of $\underline{\theta}$. Now, to obtain the joint posterior of all of the parameters, we must establish the prior distributions of the parameters.

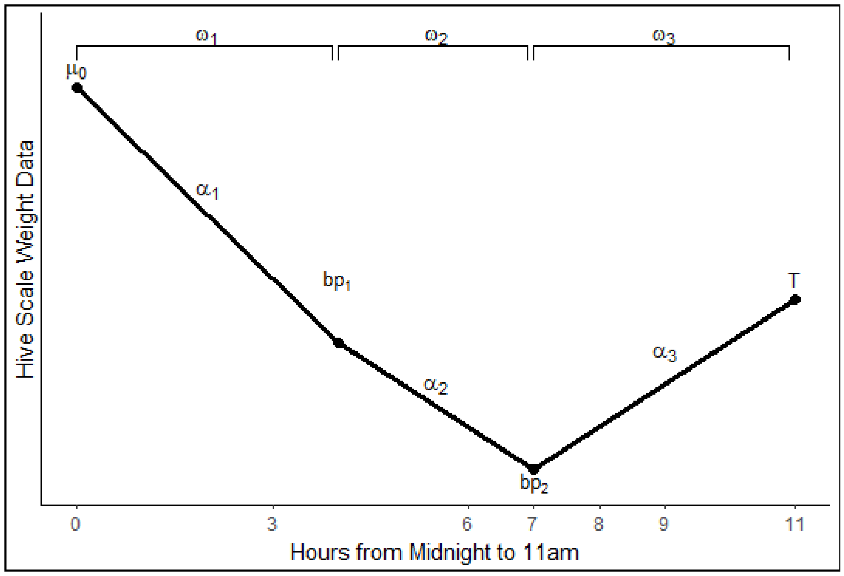

Figure 2.6: This figure gives a visualization of the parameters in the Bayesian 2 breakpoint Segmented Linear Regression model. α_1, α_2 and α_3 are the slope parameters. μ_0 is the parameter for the y-intercept. ω_1, ω_2 and ω_3 are the parameters representing the spacing between the breakpoints. bp_1 and bp_3 are the breakpoints parameters.

The regression parameters $\underline{\beta} = (\mu_0, \alpha_1, \alpha_2, \alpha_3)$, similar to the 3 breakpoint model, will have normal prior distributions with mean 0 and variance $100^2, i.e.$N(0,100^2), with the variance being sufficiently large to achieve a more non-informative prior distribution. The parameter ϵ_t is assumed to have a normal prior distribution with mean 0 and variance σ^2, where σ^2 comes from an exponential prior distribution with mean 1. To establish the prior distributions for the breakpoints, we define:

$$bp_1 = \omega_1 * T$$
$$bp_2 = (\omega_1 + \omega_2) * T$$

And specify a Dirichlet prior distribution for $\underline{\omega} = (\omega_1, \omega_2, \omega_3)'$ which is appropriate for the prior of our breakpoints as Dirichlet random variables sum to 1. Further, we put a slightly informative prior on the distribution, namely:

$$\begin{pmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{pmatrix} \sim Dir(\underline{1} * v)$$

Pattern	1st Segment	2nd Segment	3rd Segment
Pattern 1	$\hat{\alpha_1} < 0$	$\hat{\alpha_2} < 0$	$\hat{\alpha_3} > 0$
Pattern 2	$\hat{\alpha_1} > 0$	$\hat{\alpha_2} < 0$	$\hat{\alpha_3} > 0$
Pattern 3	$\hat{\alpha_1} > 0$	$\hat{\alpha_2} > 0$	$\hat{\alpha_3} < 0$
Pattern 4	$\hat{\alpha_1} < 0$	$\hat{\alpha_2} < 0$	$\hat{\alpha_3} < 0$
Pattern 5	$\hat{\alpha_1} > 0$	$\hat{\alpha_2} > 0$	$\hat{\alpha_3} > 0$
Pattern 6	$\hat{\alpha_1} < 0$	$\hat{\alpha_2} > 0$	$\hat{\alpha_3} > 0$
Pattern 7	$\hat{\alpha_1} < 0$	$\hat{\alpha_2} > 0$	$\hat{\alpha_3} < 0$
Pattern 8	$\hat{\alpha_1} > 0$	$\hat{\alpha_2} < 0$	$\hat{\alpha_3} < 0$

Table 2.3: Table the possible absolute change slope patterns possible from Meikle and Holst's definition of a Breakfast Canyon for a 2 breakpoint Segmented Linear Regression model.

So, the prior mean of $\underline{\omega}$ is $E(\underline{\omega}) = \frac{1}{\sum_{k=1}^{3} \omega_k}[\omega_1, \omega_2, \omega_3]' = \frac{1}{3}\underline{1}$. An increased v decreases the variance on the prior, making it more likely that the breakpoints are more evenly spaced. We choose $v = 3$, which results in a variance of $var(\omega_k) = \frac{\omega_k(\sum_{k=1}^{3} \omega_k - \omega_k)}{(\sum_{k=1}^{3} \omega_k)^2(\sum_{k=1}^{3} \omega_k + 1)} = \frac{1}{18}$.

2.3 Observing Breakfast Canyons in Data Following Meikle's Definition in a 2 Breakpoint SLR Model

Meikle and Holst do not explicitly define what combination of timing of breakpoints and patterns of segment slopes that correspond to a Breakfast Canyon being present in a 2 breakpoint SLR (2018). However, by utilizing similar logic to their 3 breakpoint SLR model Breakfast Canyon definition, the following definition is achieved. Again, one of the line segments with a negative slope must be followed by a segment with a positive slope. In the case of a 2 breakpoint SLR, the segment pattern can have a negative 2nd segment followed by a positive 3rd segment. Again, the segment pattern cannot be a positive 1st segment followed by a positive 2nd segment, due to the fact that ecologically the Breakfast Canyon would not commence at midnight. The possible absolute slope change patterns using Meikle and Holst's definition of a Breakfast Canyon in a 2 breakpoint Segmented Linear Regression are detailed in table 2.3

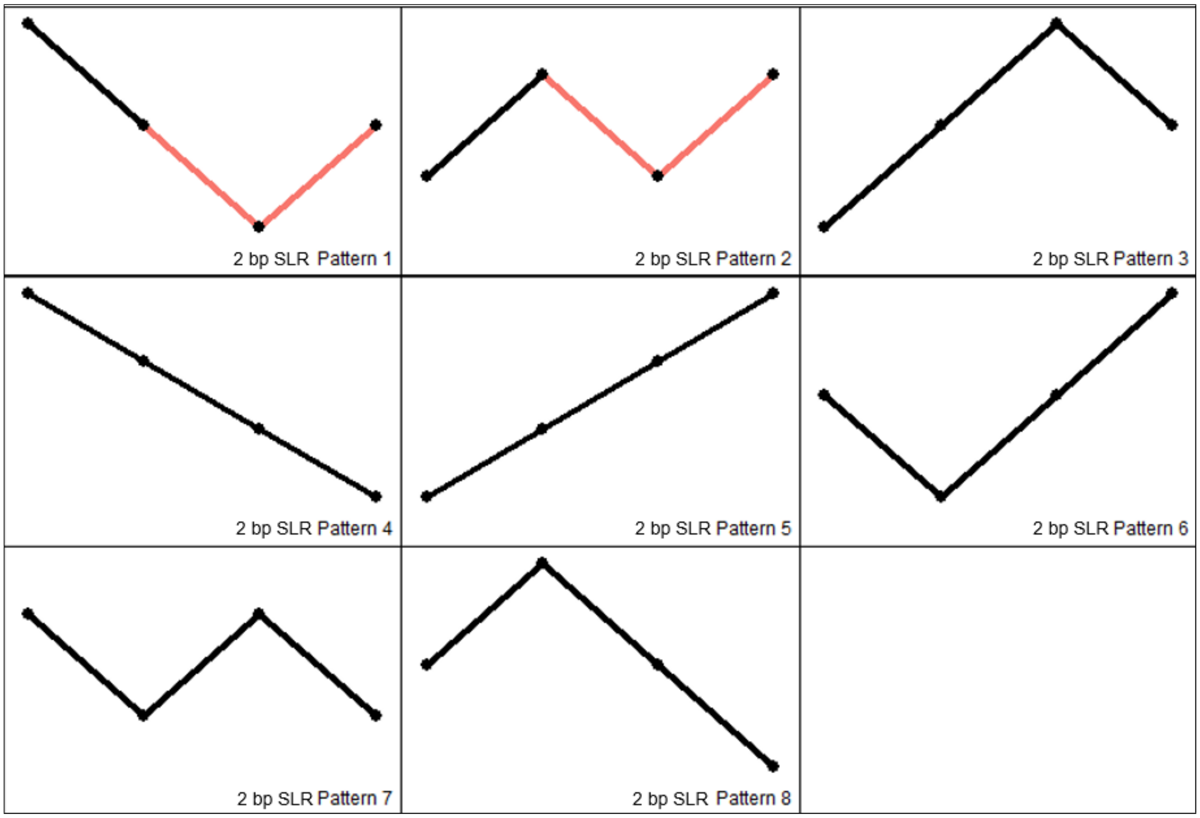

Figure 2.7: Pictured are the 8 absolute change slope patterns possible to obtain in a 2 breakpoint Segmented Linear Regression model. Patterns with segments highlighted in red are patterns that have an observed Breakfast Canyon, following the same framework as Meikle and Holst's Breakfast Canyon definition. The location of Breakfast Canyon is highlighted in red.

2.3.0.1 Slope Patterns that Result in an Observed Breakfast Canyon

In a 2 breakpoint segmented linear regression model with 3 segments, 8 absolute change slope patterns are possible, shown in figure 2.7. Of the 8 absolute change slope patterns, 2 have an observed Breakfast Canyon following Meikle's 2 breakpoint SLR model Breakfast Canyon definition, highlighted in red.

2.3.0.2 Estimating Breakfast Canyon Depth

Once we have obtained the posterior means of the slope parameters from the 2 breakpoint SLR model, notated as $\hat{\alpha_1}, \hat{\alpha_2}$, and $\hat{\alpha_3}$, as well as the posterior means of the breakpoints, notated as $\hat{bp_1}$, and $\hat{bp_2}$, we can calculate the Breakfast Canyon Depth for the SLRs in which a Breakfast Canyon was observed in the data. If a Breakfast Canyon was not observed in a hive-day case, the Breakfast Canyon Depth was considered missing.

Following Meikle's definition, the bottom of the Breakfast Canyon can only occur at the second breakpoint, and the onset of the Breakfast Canyon can only occur at the first breakpoint. We calculate the Breakfast Canyon Depth by subtracting the estimated hive weight at the second breakpoint, $\hat{bp_2}$, from the estimated hive weight at the first breakpoint , $\hat{bp_1}$. The estimated breakfast canyon depth in this case is notated as: $\hat{D} = mean_{\hat{bp_1}} - mean_{\hat{bp_2}}$

2.3.1 A Breakfast Canyon Definition Based Off of a Relative Slope Framework in a 2 Breakpoint SLR Model

One of the advantages of our relative slope definition is the ability to identify Breakfast Canyons in scenarios where Meikle's definition may not result in an observed Breakfast Canyon in the data, allowing us to observe potential Breakfast Canyons that otherwise would not be observed. Like described in section X, factors may cause an overall negative or positive weight trend that creates a barrier in observing Breakfast Canyons in data. Meikle's pattern 4, where there are three negative slopes, notated as $(\hat{\alpha_1} < 0, \hat{\alpha_2}, \hat{\alpha_3})$, 4 patterns are created under a relative slope framework. The pattern breakdown when shifting from an absolute slope framework to a relative slope framework is visualized in Figure 2.8. Again, external factors could influence the overall weight trend of the honeybee hive weight. Consider a scenario in which a dry and sunny day with low humidity follows a day of rain or high humidity. As the moisture within the hive's structure evaporates, it is expected that the hive's weight will naturally decline. Given the pre-existing downward trend in hive weight data, when the bees depart the hive for foraging, it results in a further negative slope. However, upon the bees' return from foraging, the change in weight may not be substantial enough to generate a positive slope, leading to a continuation of the negative trend, as observed in the fourth slope. Under Meikle's definition, with three negative slopes, no Breakfast Canyon would be observed. However, under our relative slope change definition, when a negative first segment slope, followed by a negative second segment slope that is more negative than the first segment slope, then followed by a third segment slope less negative than the second, a Breakfast Canyon would be observed in the data, such as Pattern 17 in Figure 2.8.

Of the 6 patterns that Meikle identified as not having an observed Breakfast Canyon,

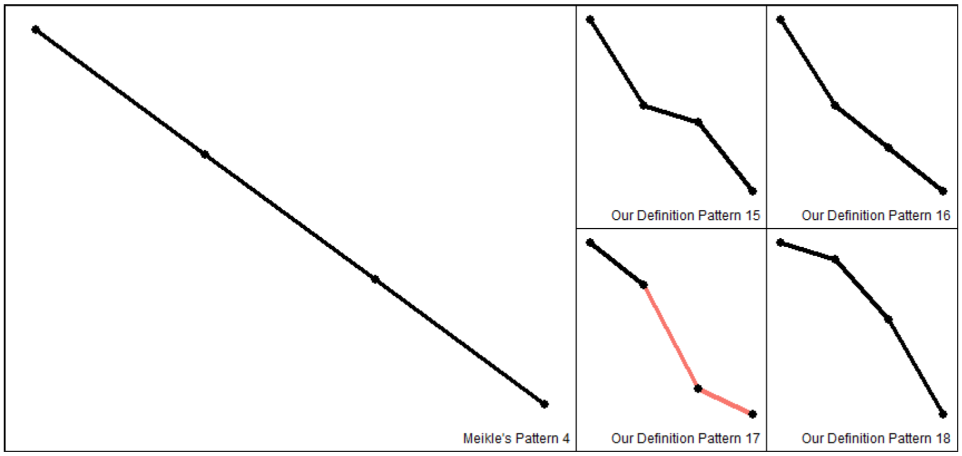

Figure 2.8: We provide a visualization of the additional patterns created from extending the Breakfast Canyon definition from Meikle's absolute slope framework to our relative slope framework, extending Pattern 4 into 4 different patterns.

15 patterns were created under our relative slope definition. Of these 15 patterns, 3 had an observed Breakfast Canyon. 18 total patterns are created under a relative slope framework for a 3 breakpoint SLR model, defined in tables 2.4.

2.3.1.1 Estimating Breakfast Canyon Depth

The process to estimate the Breakfast Canyon Depth under our relative change slope pattern Breakfast Canyon definition for a 2 breakpoint SLR model is again similar to the process defined utilizing Meikle and Holst's definition in section 2.1. For all but one pattern, the process of calculating the Breakfast Canyon depth will be the same. The pattern that will have to be approached differently is the pattern with three positive slopes. We will first detail the process of finding the Breakfast Canyon depth for the patterns that will follow the same approach as Meikle's definition. Once we have obtained the posterior means of the slope parameters from the 2 breakpoint SLR model, notated as $\hat{\alpha_1}$,$\hat{\alpha_2}$ and $\hat{\alpha_3}$, as well as the posterior means of the breakpoints, notated as $\hat{bp_1}$, and $\hat{bp_2}$, we can calculate the Breakfast Canyon Depth for the SLRs in which a Breakfast Canyon was observed in the data. If a Breakfast Canyon was not observed for a hive-day case, the Breakfast Canyon Depth was considered missing.

From our Breakfast Canyon definition for a 2 breakpoint SLR model, the bottom of the canyon can only exist at the 2nd breakpoint, and the onset of the canyon begins at

Pattern	**Segment 1**	**Segment 2**	**Segment 3**
Pattern 1	$\hat{\alpha_1} < 0$	$\hat{\alpha_2} > 0$	$\hat{\alpha_3} < 0$
Pattern 2	$\hat{\alpha_1} > 0$	$\hat{\alpha_2} < 0$	$\hat{\alpha_3} > 0$
Pattern 3	$\hat{\alpha_1} > 0$	$\hat{\alpha_2} > 0$ and $\hat{\alpha_2} \geq \hat{\alpha_1}$	$\hat{\alpha_3} < 0$
Pattern 4	$\hat{\alpha_1} > 0$	$\hat{\alpha_2} > 0$ and $\hat{\alpha_2} < \hat{\alpha_1}$	$\hat{\alpha_3} < 0$
Pattern 5	$\hat{\alpha_1} < 0$	$\hat{\alpha_2} > 0$	$\hat{\alpha_3} > 0$ and $\hat{\alpha_3} \leq \hat{\alpha_2}$
Pattern 6	$\hat{\alpha_1} < 0$	$\hat{\alpha_2} > 0$	$\hat{\alpha_3} > 0$ and $\hat{\alpha_3} > \hat{\alpha_2}$
Pattern 7	$\hat{\alpha_1} > 0$	$\hat{\alpha_2} < 0$	$\hat{\alpha_3} < 0$ and $\hat{\alpha_3} > \hat{\alpha_2}$
Pattern 8	$\hat{\alpha_1} > 0$	$\hat{\alpha_2} < 0$	$\hat{\alpha_3} < 0$ and $\hat{\alpha_3} \leq \hat{\alpha_2}$
Pattern 9	$\hat{\alpha_1} < 0$	$\hat{\alpha_2} < 0$ and $\hat{\alpha_2} < \hat{\alpha_1}$	$\hat{\alpha_3} > 0$
Pattern 10	$\hat{\alpha_1} < 0$	$\hat{\alpha_2} < 0$ and $\hat{\alpha_2} \geq \hat{\alpha_1}$	$\hat{\alpha_3} > 0$
Pattern 11	$\hat{\alpha_1} > 0$	$\hat{\alpha_2} > 0$ and $\hat{\alpha_2} < \hat{\alpha_1}$	$\hat{\alpha_3} > 0$ and $\hat{\alpha_3} > \hat{\alpha_2}$
Pattern 12	$\hat{\alpha_1} > 0$	$\hat{\alpha_2} > 0$ and $\hat{\alpha_2} < \hat{\alpha_1}$	$\hat{\alpha_3} > 0$ and $\hat{\alpha_3} \leq \hat{\alpha_2}$
Pattern 13	$\hat{\alpha_1} > 0$	$\hat{\alpha_2} > 0$ and $\hat{\alpha_2} \geq \hat{\alpha_1}$	$\hat{\alpha_3} > 0$ and $\hat{\alpha_3} > \hat{\alpha_2}$
Pattern 14	$\hat{\alpha_1} > 0$	$\hat{\alpha_2} > 0$ and $\hat{\alpha_2} \geq \hat{\alpha_1}$	$\hat{\alpha_3} > 0$ and $\hat{\alpha_3} \leq \hat{\alpha_2}$
Pattern 15	$\hat{\alpha_1} < 0$	$\hat{\alpha_2} < 0$ and $\hat{\alpha_2} \geq \hat{\alpha_1}$	$\hat{\alpha_3} < 0$ and $\hat{\alpha_3} \leq \hat{\alpha_2}$
Pattern 16	$\hat{\alpha_1} < 0$	$\hat{\alpha_2} < 0$ and $\hat{\alpha_2} \geq \hat{\alpha_1}$	$\hat{\alpha_3} < 0$ and $\hat{\alpha_3} > \hat{\alpha_2}$
Pattern 17	$\hat{\alpha_1} < 0$	$\hat{\alpha_2} < 0$ and $\hat{\alpha_2} < \hat{\alpha_1}$	$\hat{\alpha_3} < 0$ and $\hat{\alpha_3} > \hat{\alpha_2}$
Pattern 18	$\hat{\alpha_1} < 0$	$\hat{\alpha_2} < 0$ and $\hat{\alpha_2} < \hat{\alpha_1}$	$\hat{\alpha_3} < 0$ and $\hat{\alpha_3} \leq \hat{\alpha_2}$

Table 2.4: Table the possible relative change slope patterns possible from our new definition of a Breakfast Canyon for a 2 breakpoint Segmented Linear Regression model.

the 1st breakpoint. Thus, we calculate the Breakfast Canyon Depth by subtracting the estimated hive weight at the second breakpoint, $\hat{bp_2}$, from the estimated hive weight at the first breakpoint , $\hat{bp_1}$. The estimated breakfast canyon depth in this case is notated as: $\hat{D} = mean_{\hat{bp_1}} - mean_{\hat{bp_2}}$

We detail the process of calculating the Breakfast Canyon depth for the pattern with three positive slopes. We find the distance between the weight at the time at which the Breakfast Canyon reaches its' deepest point if the slope of the second segment would have been the same as the first segment, and the weight at the time the Breakfast Canyon reaches its' depth following the actual point estimate of the weight at the deepest point of the Breakfast Canyon, $mean_{\hat{bp_2}}$. To find the expected weight at the time the deepest point of the Breakfast Canyon if the second segment slope was equivalent to the first segment slope, we calculate $\hat{Ybp_2} = \hat{\alpha_1} * t_{\hat{bp_2}} + c$ where $c = mean_{\hat{bp_1}} - \hat{\alpha_1} * t_{\hat{bp_1}}$. The Breakfast Canyon depth in this case is notated as: $\hat{D} = \hat{Ybp_2} - mean_{\hat{bp_2}}$.

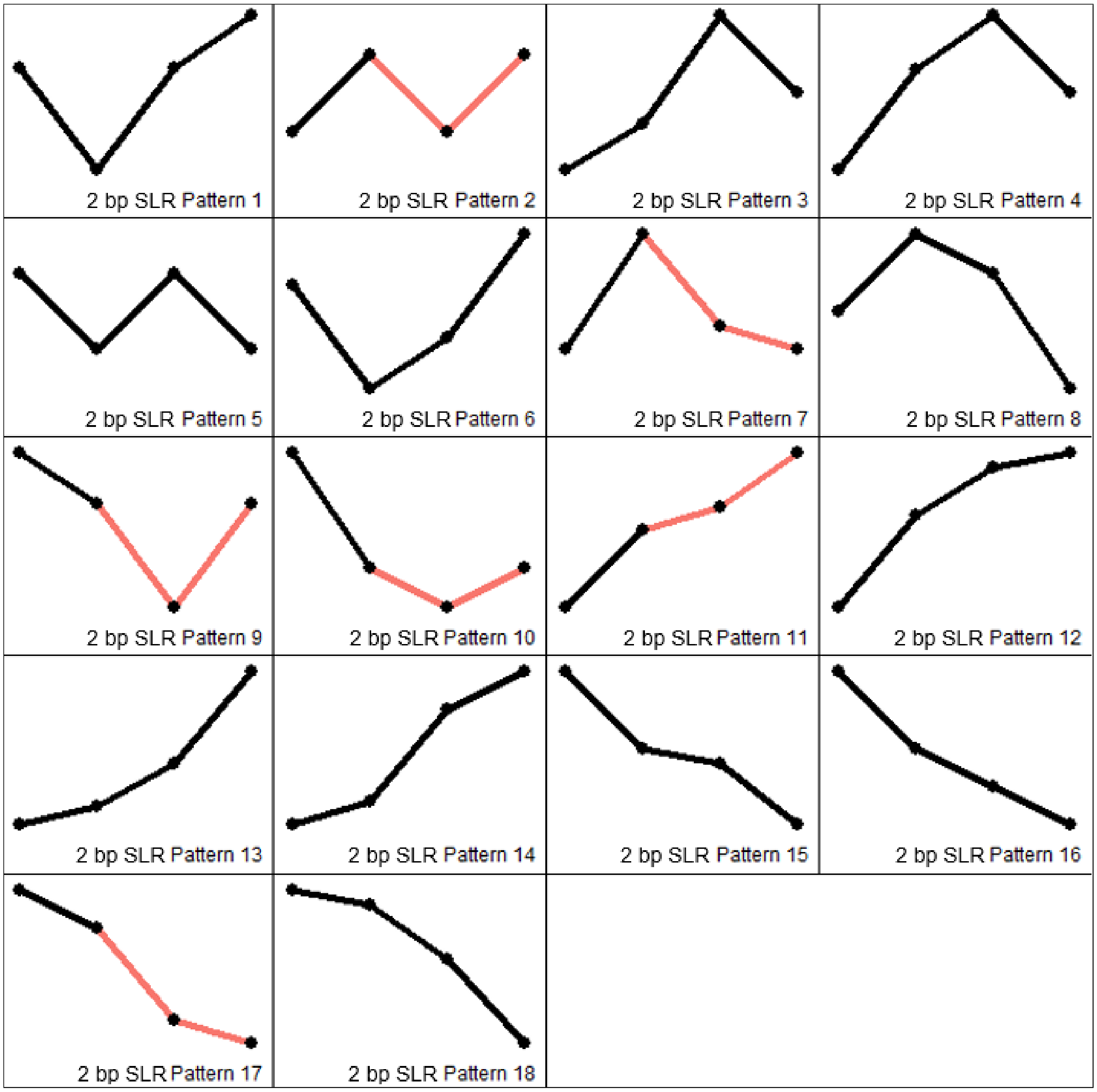

Figure 2.9: Pictured are the 18 relative change slope patterns possible to obtain in a 2 breakpoint Segmented Linear Regression model. Patterns with segments highlighted in red are patterns that have an observed Breakfast Canyon. The location of Breakfast Canyon is highlighted in red.

2.4 Model Selection and Summary Analysis

To choose between the 2 breakpoint and 3 breakpoint segmented linear regression models, we use the Wantanabe-Akaike Information Criteria (WAIC). WAIC, introduced by Watanabe (2010), is a fully Bayesian approach for estimating the out of sample log predictive likelihood (Hooten, 2015). After fitting the 2 breakpoint and 3 breakpoint segmented linear regression models for each hive-day case, we compare the models by their WAIC. We select the model by choosing the model with the smallest WAIC for that hive-day case.

After selecting the 2 or 3 breakpoint SLR model for each hive-day case, we utilize the Breakfast Canyon information for that given model. The Breakfast Canyon information includes whether or not a Breakfast Canyon was observed in the data, the location of the bottom of the Breakfast Canyon, the onset of the breakfast canyon, and the Breakfast Canyon depth. We will calculate the total proportion of days that have an observed Breakfast Canyon, as well as the number of days that follow each individual slope-pattern type for both Meikle's Breakfast Canyon definition, as well as our established Breakfast Canyon definition. Further, we will create visualizations to analyze the Breakfast Canyon depths of the days in which a Breakfast Canyon was observed in the data.

Chapter 3
Results

In this chapter, we will apply the Breakfast Canyon observation methods described in Chapter 2 to data, both real and simulated. First, we will apply the methods to real honeybee hive weight scale data from Michigan, USA. Second, we will apply the methods to simulated Brownian Motion data. For both analyses, we will provide a comparison of results between Meikle and Holst's Breakfast Canyon definition and our newly established Breakfast Canyon definition.

3.1 Example on Real Honeybee Hive Data

In this section, we will apply the methods described in Chapter 2 to real honeybee hive scale weight data, collected in Michigan. In section 3.1.1, we describe the data, and any pre-processing steps. In section 3.1.2, we describe the results of fitting the real data to our 2 and 3 breakpoint Bayesian SLR framework, the subsequent Breakfast Canyon results utilizing Meikle and Holst's definition, as well as the Breakfast Canyon results utilizing our established definition. For both BC definitions, we provide visualizations and summaries of the Breakfast Canyon information, including if a Breakfast Canyon was observed, the Breakfast Canyon Depth, which breakpoint the bottom of the Breakfast Canyon occurred, the time at which the Breakfast Canyon reaches its' bottom, what breakpoint the onset of the Breakfast Canyon occurred at, and the time at the onset of the Breakfast Canyon.

3.1.1 Description of the Data

The dataset originates from a 2019 study from the 2023 paper *Association of excessive precipitation and agricultural land use with honeybee colony performance* (Quinlan et.

al 2019), which was conducted with the objective to test if weather and landscape characteristics are associated with honeybee colony outcomes. Colony measurements, including hive weight data, were detailed and collected from over 450 honeybee colonies in Michigan, USA over the years of 2015-2018. The hive weight data was continuously collected in 15-minute time increments, like the data used in Meikle and Holst (2018).

Between 2015-2018, we have 6,712 unique hive-day observations among 450 colonies. We focused our analysis on only days in which we deemed good foraging days, defined as days with under 5mm of precipitation and over 12 degrees C. Any days that had over 5mm of precipitation or were under 12 degrees C were removed from the dataset, reducing our data from 6,712 unique hive-day instances to 5,114. Of the 5,114 remaining days, we removed days that did not include observations for the full day and days that had no weight change, indicating a potentially empty scale, reducing the number of hive-day instances to 3,930. For days that there was an impossibly large jump or dip in the weight, which may indicate a potential disruption to the hive, we pre-processed the data. We did this by replacing the impossibly large or small slope by the mean of the slope immediately before and after the jump or dip. This removed large jumps and dips in our data. After all pre-processing steps were completed, we were left with 3,930 hive-day instances.

A hypothesized Breakfast Canyon is a day-to-day occurrence, so we will analyze the individual day-hive instances independently. Further, Like Meikle and Holst (2018), we focus on the time period in which a Breafast Canyon could feasibly occur, so we only model the time period from midnight-11am, giving us 44 data points to fit our Segmented Linear Regression model to. Since we are only focusing on day-to-day weight changes, we also normalized the data by subtracting the midnight weight of each day from the weight throughout the individual day, minimizing the impact of long term weight changes from the analysis.

3.1.2 Results

In this section, we describe the results of applying our methods to the Honeybee data described above. We describe the results of the Breakfast Canyon analysis under Meikle's absolute slope Breakfast Canyon definition, as well as under our novel relative slope Breakfast Canyon definition. Further, we briefly compare the two Breakfast Canyon definition results.

3.1.2.1 Results for Real Honeybee Data Under Meikle's Absolute Slope Breakfast Canyon Definition

After applying both our 2 breakpoint and 3 breakpoint Segmented Linear Regression model framework to the 3,930 hive-day instances, we conducted model selection by comparing WAIC between the two models, selecting the Model with the lower WAIC. We use the Nimble package in R to generate and execute our MCMC algorithm, with 100,000 iterations and 1,000 burn-in. For each of the 3,930 instances, it takes approximately 42 seconds to run our MCMC algorithm under the 2 breakpoint SLR framework in nimble, thus a total time of 45 hours and 51 minutes. Under the 3 breakpoint SLR framework in nimble, it takes approximately 44 seconds to run our MCMC algorithm for each of the 3,930 instances, taking a total time of 48 hours and 2 minutes. We execute our MCMC algorithm on an intel i7 processor with 4.70 GHz. Confirmation of convergence of the MCMC algorithm was completed for most days through visual inspection of the MCMC trace plots of the parameters. For each hive-day instance, we choose the model that has the lower WAIC. Of the 3,930 hive-day instances, the 2 breakpoint Segmented Linear Regression model was chosen 573 times, and the 3 breakpoint Segmented Linear Regression model was chosen 3,357 times. After selecting a 3 or 2 breakpoint Segmented Linear Regression model for each of the hive-day instances, we applied Meikle's corresponding 2 or 3 breakpoint Breakfast Canyon definition to observe whether or not a Breakfast Canyon exists in the data, and collect the Breakfast Canyon information. We also applied our novel relative slope 2 and 3 breakpoint Breakfast Canyon definition to observe whether or not a Breakfast Canyon exists in the data, and again collect the Breakfast Canyon information.

Under Meikle's absolute slope Breakfast Canyon definition, 207 of the 3,930 hive-day instances had an observed Breakfast Canyon. Thus, 5.27% of the hive-day instances had posterior mean slope patterns which indicated a Breakfast Canyon was present. Of these 207 hive-day instances, 36 of the hive-day had a chosen 2 breakpoint Segmented Linear Regression model, while the remaining 171 had a chosen 3 breakpoint Segmented Linear Regression Model.

The 207 hive-day instances where a Breakfast Canyon was observed fall into 8 of the 12 possible Breakfast Canyon patterns between the 2 breakpoint and 3 breakpoint absolute slope patterns that result in an observed Breakfast Canyon in the data under Meikle's absolute slope framework. The patterns that the hive-day instances do not

Model Selected	**Pattern**	**Number of Instances**
3 breakpoint	Pattern 8	71
3 breakpoint	Pattern 1	41
3 breakpoint	Pattern 17	24
2 breakpoint	Pattern 1	23
3 breakpoint	Pattern 5	23
2 breakpoint	Pattern 2	12
3 breakpoint	Pattern 4	8
3 breakpoint	Pattern 18	4

Table 3.1: Described in the table above are the number of hive-day instances that fall into the 8 observed Breakfast Canyon patterns in our Honeybee data under Meikle's absolute slope Breakfast Canyon definition.

fall into include patterns 2, 3, 6 and 7 from a 3 breakpoint absolute slope framework. A numerical summary of the Breakfast Canyon patterns observed under a absolute slope framework are shown in Table 3.1. The most common Breakfast Canyon pattern observed in the Honeybee data under an absolute slope framework was Pattern 8 for a 3 breakpoint SLR model, with 71 hive-day instances falling into the pattern. After Pattern 8, Patterns 1 and 17 from a 3 breakpoint SLR model were most commonly observed in the Honeybee data under Meikle's absolute slope framework. Then, Pattern 5 from a 3 breakpoint SLR framework and patterns 1 and 2 from a 2 breakpoint SLR framework were commonly found. Posterior mean plots for one hive-day instance following each pattern for the top 5 observed Breakfast Canyon patterns can be shown in Figure 3.1 .

The Breakfast Canyon reached its' depth at the 2nd breakpoint 35 of the hive-day instances. All of the 35 hive-day instances hive-day instances where a 2 breakpoint SLR model was chosen. The Breakfast Canyon reached its' depth at the 3rd breakpoint in 171 of the hive-day instances, all of which could only be instances in which a 3 breakpoint SLR model was chosen. A summary of the breakpoint location of the canyon depth for days with an observed Breakfast Canyon can be found in Table 3.2.
None of the day-hive instances with the onset beginning at the first breakpoint reached the bottom of the canyon at the 3rd breakpoint, meaning that the Breakfast Canyon was comprised of the 2nd, 3rd, and 4th segments.

The mean Breakfast Canyon depth for the Honeybee data under Meikle's absolute slope framework is 0.157 kg, with a standard deviation of 0.204 kg. The median Breakfast

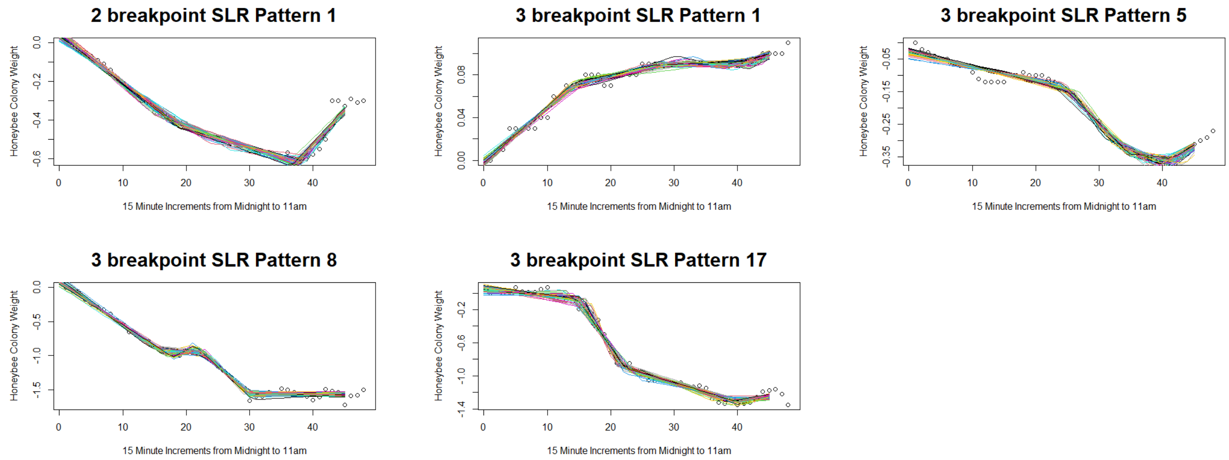

Figure 3.1: For the top 5 patterns with an observed Breakfast Canyon using Meikle's absolute slope Breakfast Canyon definition, we chose one hive-day instance per pattern. We plot the raw data for the hive-day instance, shown as the points on the plot. Further, we plot 100 samples from the posterior distribution of a mean hive weight, which are the colored lines of the plot.

Model Selected	**Location of the Bottom of the Breakfast Canyon**		**Location of the Onset of the Breakfast Canyon**	
	2nd Breakpoint	*3rd Breakpoint*	*1st Breakpoint*	*2nd Breakpoint*
2 breakpoint SLR model	35		35	
3 breakpoint SLR model	0	171	0	171

Table 3.2: A summary of the counts of the locations of the bottom of the Breakfast Canyon under Meikle's Breakfast Canyon definition, as well as the locations of the onset of the Breakfast Canyon. Counts are broken down by the model that was selected via WAIC, which is either the 2 or 3 breakpoint Segmented Linear Regression Model.

Canyon depth for the Honeybee data is 0.091 kg, and the distribution of the Breakfast Canyon Depth is right skewed, shown in Figure 3.2.

The mean time for Breakfast Canyon's deepest point for the Honeybee data under Meikle's absolute slope framework is around 7:25 AM, with a standard deviation of 1 hour 14 minutes, and a median timing of the Breakfast Canyon's deepest point of around 7:24 AM. A visualization of the distribution of the Breakfast Canyon deepest point timing is shown in Figure 3.3.

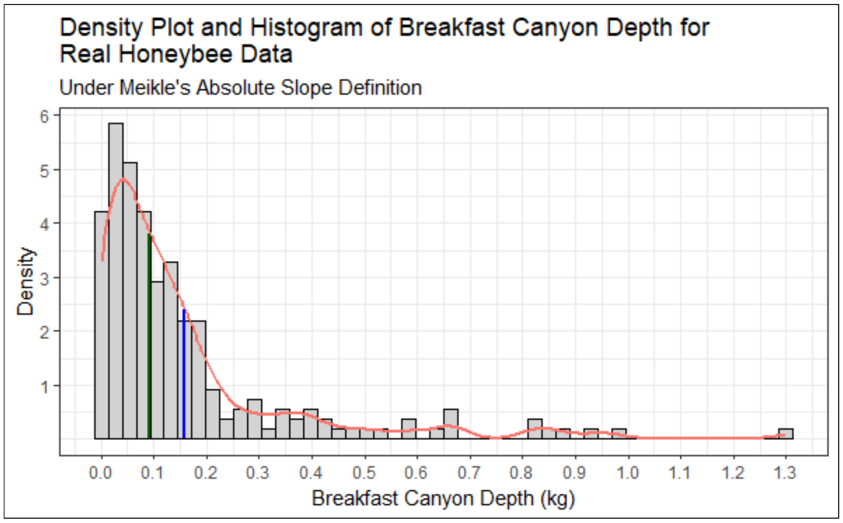

Figure 3.2: Pictured is a histogram of the distribution of the Breakfast Canyon Depth for the Real Honeybee data, using Meikle's absolute slope definition for Breakfast Canyon observation, with the density curve plotted in red. The median of the Breakfast Canyon Depth is the green line, while the mean of the Breakfast Canyon Depth is the blue line.

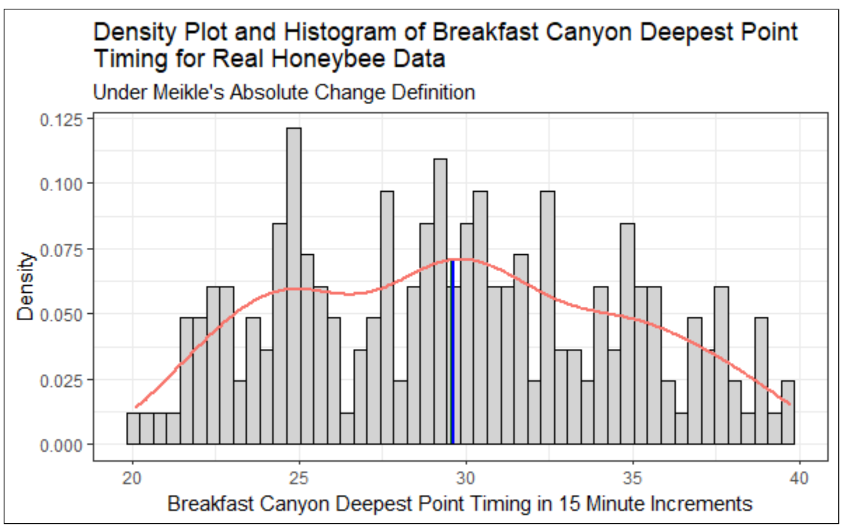

Figure 3.3: Pictured is a histogram of the distribution of the time of the deepest point of the Breakfast Canyon for the Real Honeybee data, using Meikle's absolute slope definition for Breakfast Canyon observation, with the density curve plotted in red. The median of the time of the deepest point is the green line, while the mean of the time of the deepest point is the blue line.

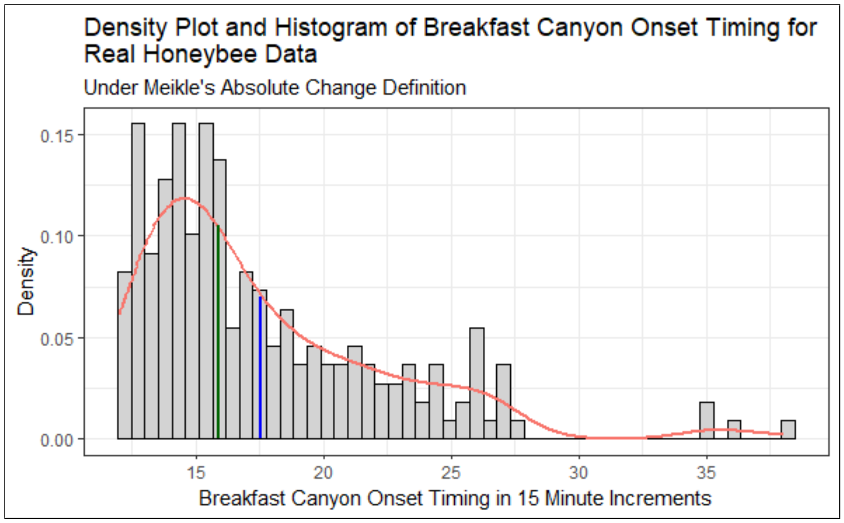

Figure 3.4: Pictured is a histogram of the distribution of the time of the onset of the Breakfast Canyon for the Real Honeybee data, using Meikle's absolute slope definition for Breakfast Canyon observation, with the density curve plotted in red. The median of the time of the Breakfast Canyon onset is the green line, while the mean of the time of the Breakfast Canyon onset is the blue line.

The onset of the Breakfast Canyon begins at the 1st breakpoint 35 of the hive-day instances. All of the 35 hive-day instances were instances where a 2 breakpoint SLR model was chosen. The onset of the Breakfast Canyon began at the 2nd breakpoint 171 of the hive-day instances, and all of which were only instances in which a 3 breakpoint SLR model was chosen. A summary of the location of the onset of the Breakfast Canyon is provided in Table 3.2.

The mean time for Breakfast Canyon onset for the Honeybee data under Meikle's absolute slope framework is around 4:23 AM, a standard deviation of 1 hour 12 minutes, and a median Breakfast Canyon onset of around 3:58 AM. A visualization of the distribution of the Breakfast Canyon Onset timing is shown in Figure 3.4.

Lastly, we plot the extracted line segments for the observed Breakfast Canyon days from the onset of the Breakfast Canyon to the deepest point of the canyon. We normalized the data to have all segments begin at zero, thus allowing us to visualize the depth of each Breakfast Canyon hive-day instance, by Pattern type. The breakfast Canyon Depths are visualized in Figure 3.5

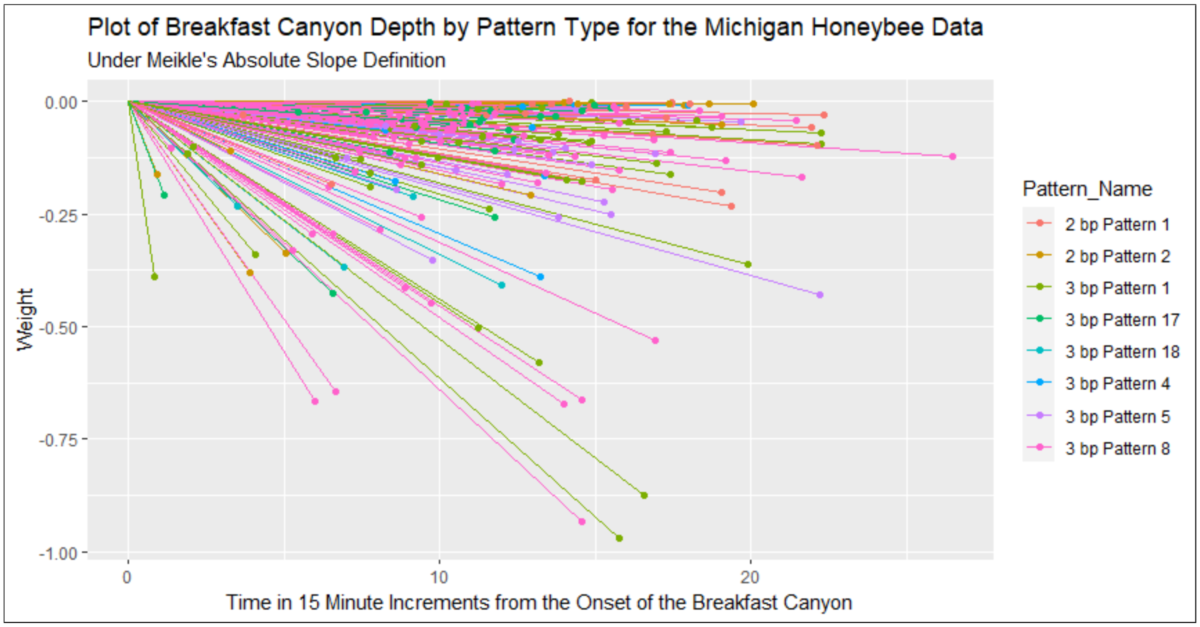

Figure 3.5: Pictured are the extracted line segments used to estimate the Breakfast Canyon depth in kg for days with an observed Breakfast Canyon (n=207). The segments were normalized to begin at zero weight, with time at zero being the time of the onset of the Breakfast Canyon. Line segments and endpoints were color-coded by Pattern Type.

3.1.2.2 Results for Real Honeybee Data Under Our Relative Slope Breakfast Canyon Definition

Under our novel relative slope Breakfast Canyon definition, 347 of the 3,930 hive-day instances had an observed Breakfast Canyon. Thus, 8.83% of the hive-day instances had posterior mean slope patterns which indicated a Breakfast Canyon was present. Of these 347 hive-day instances, 47 of the hive-day had a chosen 2 breakpoint Segmented Linear Regression model, while the remaining 300 had a chosen 3 breakpoint Segmented Linear Regression Model.

The 347 hive-day instances where a Breakfast Canyon was observed fall into 26 of the 41 possible Breakfast Canyon patterns between the 2 breakpoint and 3 breakpoint relative slope patterns that result in an observed Breakfast Canyon. A numerical summary of the Breakfast Canyon patterns observed under a relative slope framework are shown in Table 3.3. In the 2 breakpoint SLR model framework, only one Breakfast Canyon pattern was not observed in the Honeybee data, which was Pattern 11. In the 3 breakpoint SLR model framework, 14 Breakfast Canyon patterns were not observed in the real Honeybee data, which include Patterns 1, 2, 3, 4, 5, 6, 12, 28, 33, 41, 45, 46, 51 and 53. Thus, a total of

Model Selected	**Pattern**	**Number of Instances**
3 breakpoint	Pattern 17	71
3 breakpoint	Pattern 48	25
3 breakpoint	Pattern 39	24
3 breakpoint	Pattern 20	21
3 breakpoint	Pattern 7	19
3 breakpoint	Pattern 15	19
3 breakpoint	Pattern 40	17
2 breakpoint	Pattern 10	17
3 breakpoint	Pattern 25	14
3 breakpoint	Pattern 18	13
3 breakpoint	Pattern 10	12
2 breakpoint	Pattern 2	12
3 breakpoint	Pattern 21	11
3 breakpoint	Pattern 31	10
3 breakpoint	Pattern 30	8
3 breakpoint	Pattern 38	7
2 breakpoint	Pattern 7	7
2 breakpoint	Pattern 9	7
3 breakpoint	Pattern 14	6
3 breakpoint	Pattern 36	6
3 breakpoint	Pattern 54	6
3 breakpoint	Pattern 11	5
2 breakpoint	Pattern 17	4
3 breakpoint	Pattern 34	3
3 breakpoint	Pattern 19	2
3 breakpoint	Pattern 29	1

Table 3.3: Described in the table above are the number of hive-day instances that fall into the 26 observed Breakfast Canyon patterns in our Honeybee data under our relative slope Breakfast Canyon definition.

15 relative slope patterns with an observed Breakfast canyon were not present in the real Honeybee data. The most common Breakfast Canyon pattern observed in the Honeybee data was Pattern 17 for a 3 breakpoint SLR model, yielding 71 hive-day instances that fell into this pattern. After Pattern 17, Patterns 48, 39, 20, 7, 15, and 14 from a 3 breakpoint SLR model, and Pattern 10 from a 2 breakpoint SLR model were most commonly observed in the real Honeybee data. Posterior mean plots for one hive-day instance following each pattern for the top 5 observed Breakfast Canyon patterns can be shown in Figure 3.6.

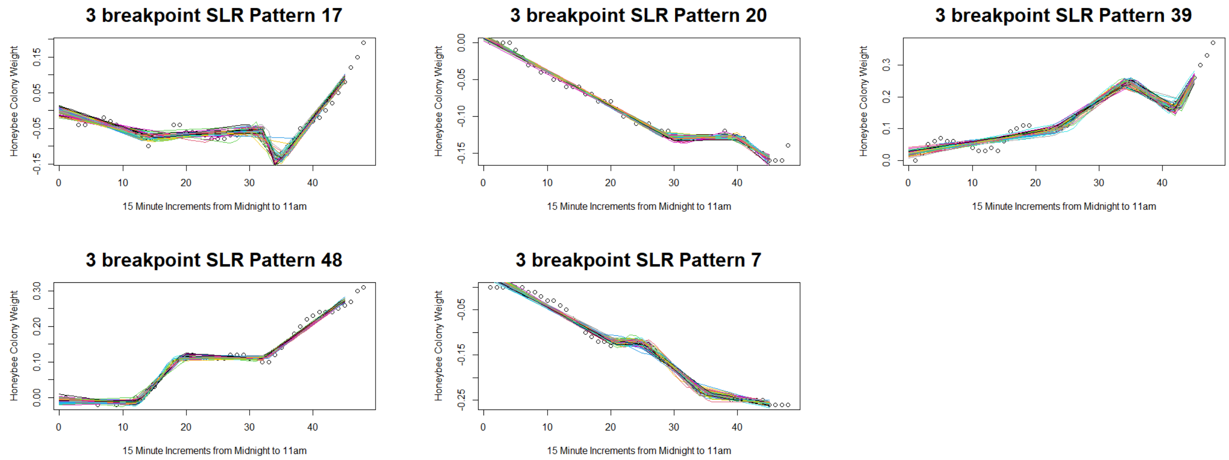

Figure 3.6: For the top 5 patterns with an observed Breakfast Canyon using our relative slope Breakfast Canyon definition, we chose one hive-day instance per pattern. We plot the raw data for the hive-day instance, shown as the points on the plot. Further, we plot 100 means from the posterior distribution, which are the colored lines of the plot.

Model Selected	**Location of the Bottom of the Breakfast Canyon**		**Location of the Onset of the Breakfast Canyon**	
	2nd Breakpoint	*3rd Breakpoint*	*1st Breakpoint*	*2nd Breakpoint*
2 breakpoint SLR model	47		47	
3 breakpoint SLR model	0	300	0	474

Table 3.4: A summary of the counts of the locations of the bottom of the Breakfast Canyon under our relative slope Breakfast Canyon definition, as well as the locations of the onset of the Breakfast Canyon. Counts are broken down by the model that was selected via WAIC, which is either the 2 or 3 breakpoint Segmented Linear Regression Model.

The Breakfast Canyon reached its' depth at the 2nd breakpoint 47 of the hive-day instances. All of the 47 instances were hive-day instances where a 2 breakpoint SLR model was chosen. The Breakfast Canyon reached its' depth at the 3rd breakpoint 300 of the hive-day instances, all of which could only be instances in which a 3 breakpoint SLR model was chosen. A summary of the breakpoint location of the canyon depth for days with an observed Breakfast Canyon can be found in Table 3.4.
None of the day-hive instances with the onset beginning at the first breakpoint reached the bottom of the canyon at the 3rd breakpoint, meaning that the Breakfast Canyon was comprised of the 2nd, 3rd, and 4th segments.

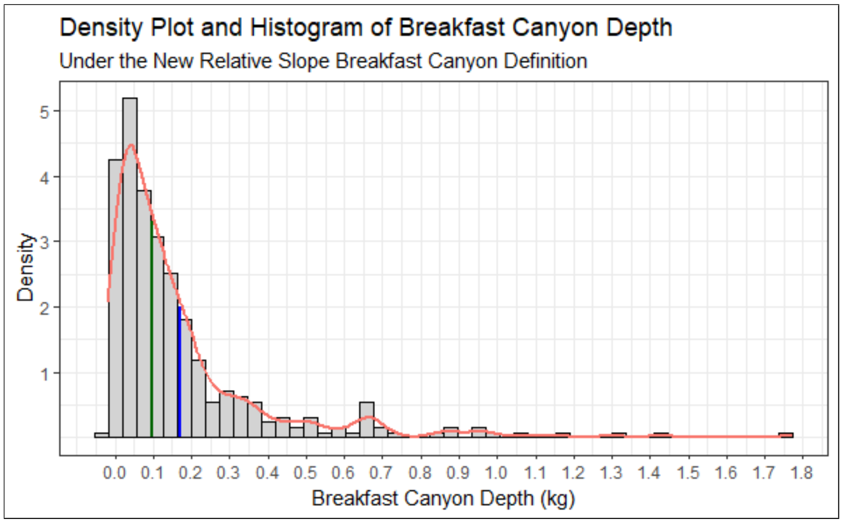

Figure 3.7: Pictured is a histogram of the distribution of the Breakfast Canyon Depth for the Real Honeybee data, using our relative slope definition for Breakfast Canyon observation, with the density curve plotted in red. The median of the Breakfast Canyon Depth is the green line, while the mean of the Breakfast Canyon Depth is the blue line.

The mean Breakfast Canyon depth for the Honeybee data under our relative slope definition framework is 0.171 kg, with a standard deviation of 0.233 kg. The median Breakfast Canyon depth for the Honeybee data is 0.097 kg. A visualization of the distribution of the Breakfast Canyon depth is shown in Figure 3.7.

The mean time for Breakfast Canyon's deepest point for the Honeybee data is about 7:10 AM, with a standard deviation of about 1 hour 13 minutes, and a median timing of the Breakfast Canyon's deepest point of around 7:03 AM. A visualization of the distribution of the Breakfast Canyon deepest point timing is shown in Figure 3.8.

The onset of the Breakfast Canyon begins at the 1st breakpoint 47 of the hive-day instances. All of the 47 hive-day instances were instances where a 2 breakpoint SLR model was chosen. The onset of the Breakfast Canyon began at the 2nd breakpoint 300 of the hive-day instances, and all of which were only instances in which a 3 breakpoint SLR model was chosen. A summary of the location of the onset of the Breakfast Canyon is provided in Table 3.4.

The mean time for Breakfast Canyon onset for the Honeybee data under our relative slope framework is around 4:20 AM, a standard deviation of 1 hour 9 minutes, and a median Breakfast Canyon onset of around 3:59 AM. A visualization of the distribution

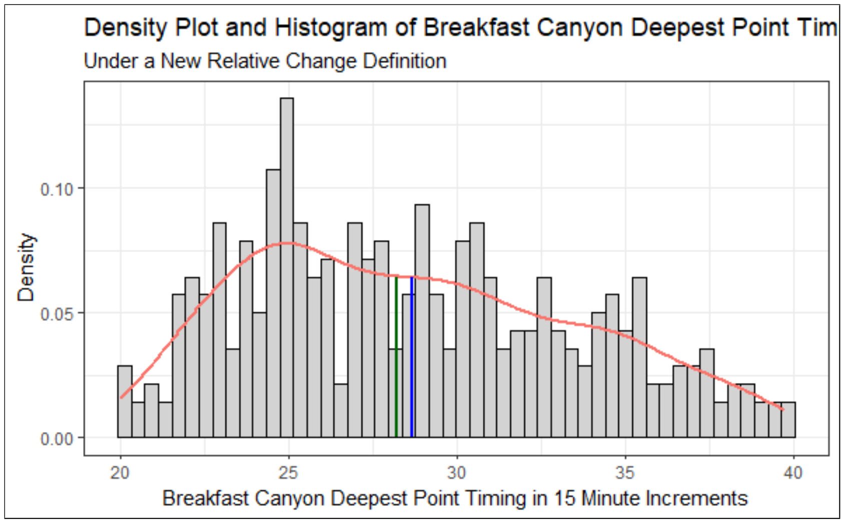

Figure 3.8: Pictured is a histogram of the distribution of the time of the deepest point of the Breakfast Canyon for the Real Honeybee data, using our relative slope definition for Breakfast Canyon observation, with the density curve plotted in red. The median of the time of the deepest point is the green line, while the mean of the time of the deepest point is the blue line.

of the Breakfast Canyon Onset timing is shown in Figure 3.9.

Lastly, we plot the extracted line segments for the observed Breakfast Canyon days from the onset of the Breakfast Canyon to the deepest point of the canyon. We normalized the data to have all segments begin at zero, allowing us to visualize the depth of each Breakfast Canyon hive-day instance, by Pattern type, visualized in Figure 3.10

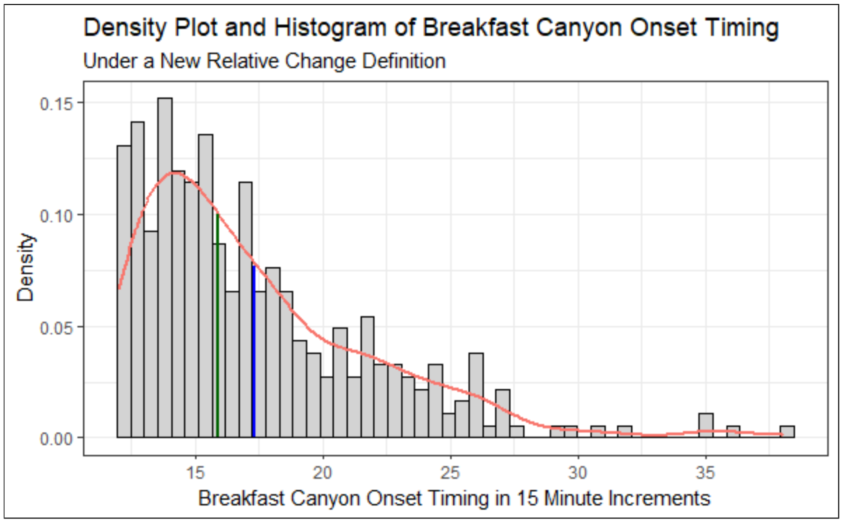

Figure 3.9: Pictured is a histogram of the distribution of the time of the onset of the Breakfast Canyon for the Real Honeybee data, using our relative slope definition for Breakfast Canyon observation, with the density curve plotted in red. The median of the time of the Breakfast Canyon onset is the green line, while the mean of the time of the Breakfast Canyon onset is the blue line.

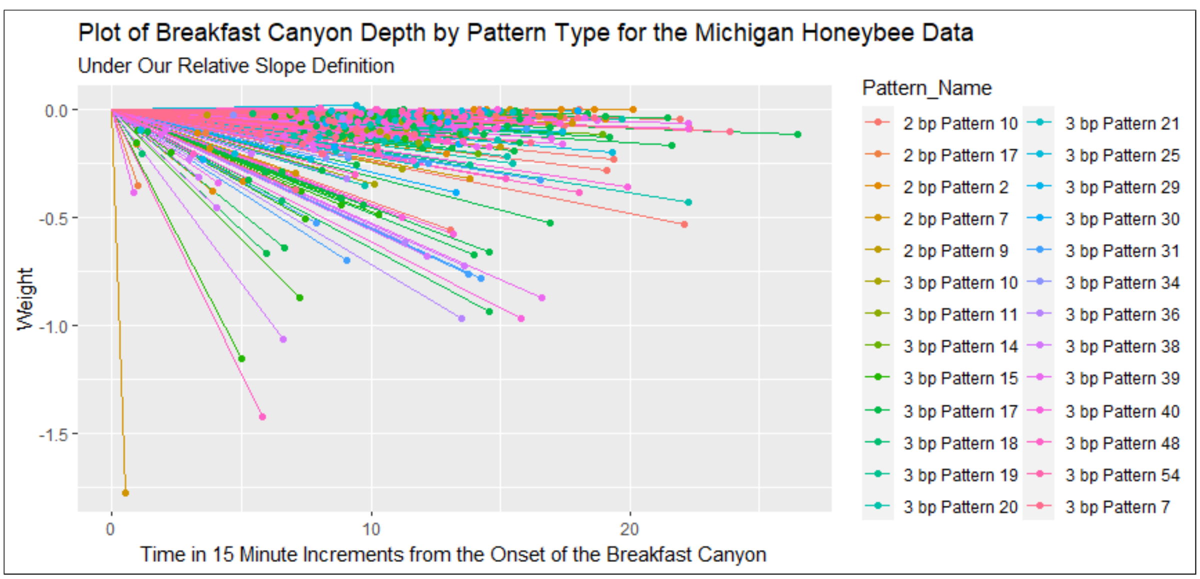

Figure 3.10: Pictured are the extracted line segments used to estimate the Breakfast Canyon depth in kg for days with a observed Breakfast Canyon (n=347). The segments were normalized to begin at zero weight, with time at zero being the time of the onset of the Breakfast Canyon. Line segments and endpoints were color-coded by Pattern Type.

3.1.2.3 A Comparison of the Results Between Breakfast Canyon Definitions for Real Honeybee Data

We next compare the results of applying Meikle's absolute slope definition and Our new relative slope definition to Honeybee data. We first observe that Meikle's absolute slope definition yields a proportion of 5.27% of days with an observed Breakfast Canyon, while Our relative slope definition yields a proportion of 8.83% of days with an observed Breakfast Canyon. Since our patterns extend upon Meikle's patterns, a larger proportion of observed Breakfast Canyon days is what we would expect to see theoretically.

We observe when comparing the distributions of the Breakfast Canyon Depth in Kilograms between the two definitions (Figures 3.7 and 3.2) that they both result in a right-skewed distributions. Similarly, the distributions of the timing of the deepest point of the Breakfast Canyon are visually similar when comparing between the two definitions (Figures 3.8 and 3.3). Further, the distributions of the timing of the onset of the Breakfast Canyon are visually similar when comparing the two definitions (Figures 3.9 and 3.4). Further, when comparing the extracted segments that are used to estimate Breakfast Canyon depth, shown in Figures 3.10 and 3.5, some of the extracted line segments have a shown positive slope under Our relative definition, while none of the extracted line segments have a shown negative slope under our relative definition, due to the extension of our definition.

3.2 Example on Simulated Brownian Motion Data

In this section, we will apply the methods described in Chapter 2 to simulation data, created from simulated Brownian Motion. In section 3.2.1, we describe the process of simulating the data from Brownian motion. In section 3.2.2, we describe the results of fitting the simulated data to our 2 and 3 breakpoint Bayesian SLR framework, the subsequent Breakfast Canyon results utilizing Meikle and Holst's definition, as well as the Breakfast Canyon results utilizing our established definition. As we did for the real data in section 3.1, for both BC definitions, we provide visualizations and summaries of the Breakfast Canyon information, including if a Breakfast Canyon was observed, the Breakfast Canyon Depth, what breakpoint and time the bottom of the Breakfast Canyon occurred, and what breakpoint and time the onset of the Breakfast Canyon occurred. The goal of applying our framework to simulated Brownian Motion is to compare the simulated data to the real data to explore the existence of the Breakfast Canyon in real data when compared to random noise.

3.2.1 Data Simulated from Brownian Motion

Brownian motion is a stochastic model, in which the increments (changes from one time point to the next) are distributed as a normal distribution with mean 0 and variance $\sigma^2 * \Delta t$ (Schilling et. al., 2021). We simulated 1,000 instances of Brownian Motion for 44 discrete timepoints in which we specify the variance of the diffusion process as $\sigma^2 = 0.4$. This process creates 1,000 simulated hive-day instances, in which each instance has 44 timepoints, representings the 15-minute increments present between midnight and 11am. One simulated Brownian Motion instance, with 44 timepoints, is visualized in figure 3.11. The 1,000 simulated Brownian Motion instances are visualized in figure 3.12, plotted together on one graph.

3.2.2 Results

In this section, we describe the results of applying our methods to simulated Brownian Motion data. We describe the results of the Breakfast Canyon analysis under Meikle's absolute slope Breakfast Canyon definition, as well as under Our novel relative slope Breakfast Canyon definition. Further, we briefly compare the two Breakfast Canyon definition results.

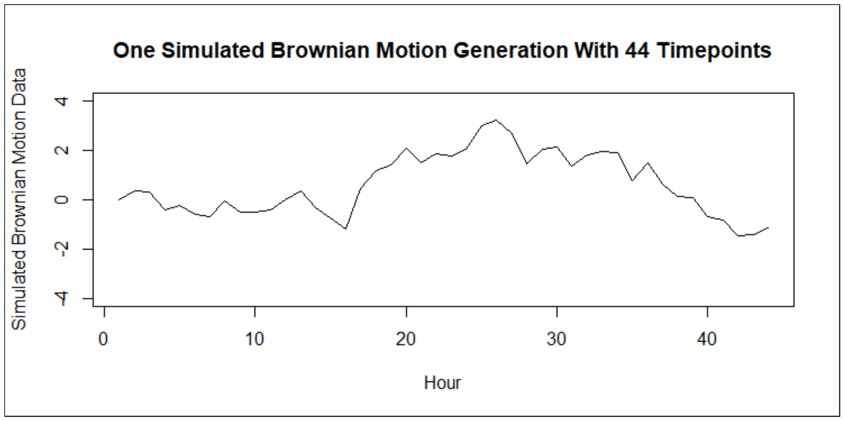

Figure 3.11: Pictured is one of the simulated Brownian Motion instances with 44 timepoints.

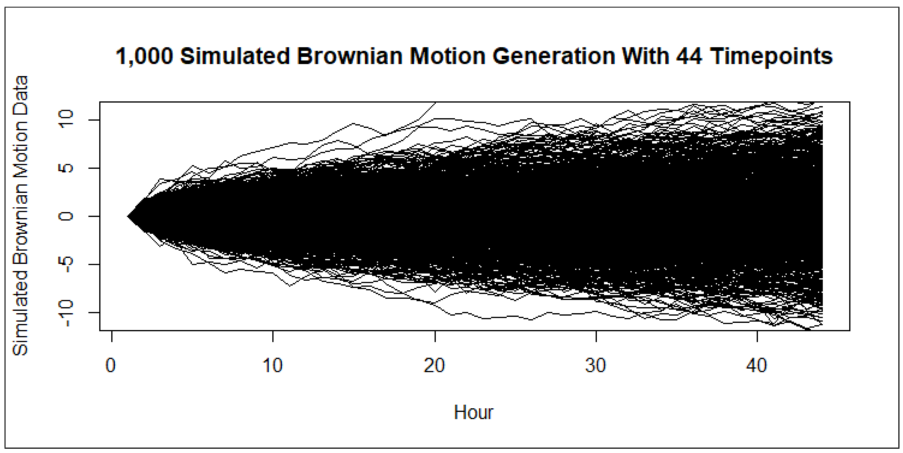

Figure 3.12: Pictured are the 1,000 simulated Brownian Motion instances with 44 timepoints, plotted together.

3.2.2.1 Results for Brownian Motion Under Meikle's Absolute Slope Breakfast Canyon Definition

Under Meikle's absolute slope Breakfast Canyon definition, 165 of the 1,000 hive-day instances had an observed Breakfast Canyon. Thus, 16.5% of the hive-day instances had posterior mean slope patterns which indicated a Breakfast Canyon was present. Of these 165 hive-day instances, 38 of the hive-day had a chosen 2 breakpoint Segmented Linear Regression model, while the remaining 127 had a chosen 3 breakpoint Segmented Linear Regression Model. We use the Nimble package in R to generate and execute our MCMC algorithm, with 100,000 iterations and 1,000 burn-in. For each of the 1,000 instances, it takes approximately 42 seconds to run our MCMC algorithm under the 2 breakpoint SLR framework in nimble, thus a total time of 11 hours and 40 minutes. Under the

Model Selected	Pattern	Number of Instances
3 breakpoint	Pattern 8	67
2 breakpoint	Pattern 2	29
3 breakpoint	Pattern 1	28
3 breakpoint	Pattern 4	11
2 breakpoint	Pattern 1	9
3 breakpoint	Pattern 5	7
3 breakpoint	Pattern 17	7
3 breakpoint	Pattern 18	7

Table 3.5: Described in the table above are the number of simulated Brownian Motion instances that fall into the 8 observed Breakfast Canyon patterns in our simulated Brownian Motion data under Meikle's absolute slope Breakfast Canyon definition.

3 breakpoint SLR framework in nimble, it takes approximately 44 seconds to run our MCMC algorithm for each of the 1,000 instances, taking a total time of 12 hours and 13 minutes. We execute our MCMC algorithm on an intel i7 processor with 4.70 GHz. Convergence of parameters was completed for most days through visual inspection of the trace plots.

The 165 hive-day instances where a Breakfast Canyon was observed fall into 8 of the 12 possible Breakfast Canyon patterns between the 2 breakpoint and 3 breakpoint absolute slope patterns that correspond in a Breakfast Canyon being present in the data. The absolute slope Breakfast Canyon patterns not observed in the simulated Brownian Motion data include patterns 2, 3, 6 and 7 under a 3 breakpoint SLR framework. A numerical summary of the Breakfast Canyon patterns observed under a absolute slope framework are shown in Table 3.5. The most common Breakfast Canyon pattern observed in the Brownian Motion data under an absolute slope framework was Pattern 8 for a 3 breakpoint SLR model, with 67 hive-day instances falling into the pattern. Following Pattern 8, Pattern 2 from a 2 breakpoint SLR model, and Pattern 1 from a 3 breakpoint SLR model were most commonly observed in the Brownian Motion data. Following those patterns, patterns 4 and 5 from a 3 breakpoint SLR framework were commonly observed, along with pattern 1 from a 2 breakpoint SLR framework. Posterior mean plots for one hive-day instance following each pattern for the top 5 observed patterns can be shown in Figure 3.13.

The Breakfast Canyon reached its' depth at the 2nd breakpoint 38 of the hive-day

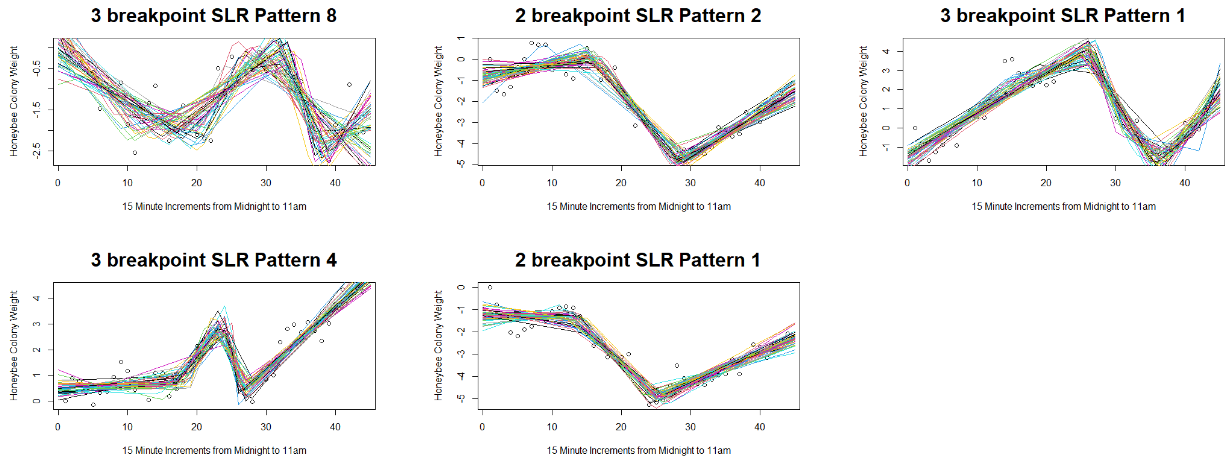

Figure 3.13: For the top 5 patterns with an observed Breakfast Canyon using Meikle's absolute slope Breakfast Canyon definition, we chose one simulated brownian motion instance per pattern. We plot the raw simulated brownian motion instance, shown as the points on the plot. Further, we plot 100 means from the posterior distribution, which are the colored lines of the plot.

instances. All of the 38 hive-day instances were instances where a 2 breakpoint SLR model was chosen. The Breakfast Canyon reached its' depth at the 3rd breakpoint 127 of the hive-day instances, all of which could only be instances in which a 3 breakpoint SLR model was chosen. A summary of the breakpoint location of the canyon depth for days with an observed Breakfast Canyon can be found in Table 3.6. None of the day-hive instances with the onset beginning at the first breakpoint reached the bottom of the canyon at the 3rd breakpoint, meaning that the Breakfast Canyon was comprised of the 2nd, 3rd, and 4th segments.

The mean Breakfast Canyon depth for the Brownian Motion data under Meikle's absolute slope framework is 3.092 kg, with a standard deviation of 1.545 kg. The median Breakfast Canyon depth for the Brownian Motion data is 3.056 kg, and the distribution of the Breakfast Canyon Depth is roughly symmetric, shown in Figure 3.14.

The mean time for Breakfast Canyon's deepest point for the Brownian Motion data under Meikle's absolute slope framework is around 6:54 AM, a standard deviation of 1 hour 9 minutes, and a median timing of the Breakfast Canyon's deepest point of around 6:41 AM. A visualization of the distribution of the Breakfast Canyon deepest point timing is shown in Figure 3.15.

Model Selected	**Location of the Bottom of the Breakfast Canyon**		**Location of the Onset of the Breakfast Canyon**	
	2nd Breakpoint	*3rd Breakpoint*	*1st Breakpoint*	*2nd Breakpoint*
2 breakpoint SLR model	38		38	
3 breakpoint SLR model	0	127	0	127

Table 3.6: A summary of the counts of the locations of the bottom of the Breakfast Canyon for Brownian Motion simulated data under Meikle's absolute slope Breakfast Canyon definition, as well as the locations of the onset of the Breakfast Canyon. Counts are broken down by the model that was selected via WAIC, which is either the 2 or 3 breakpoint Segmented Linear Regression Model.

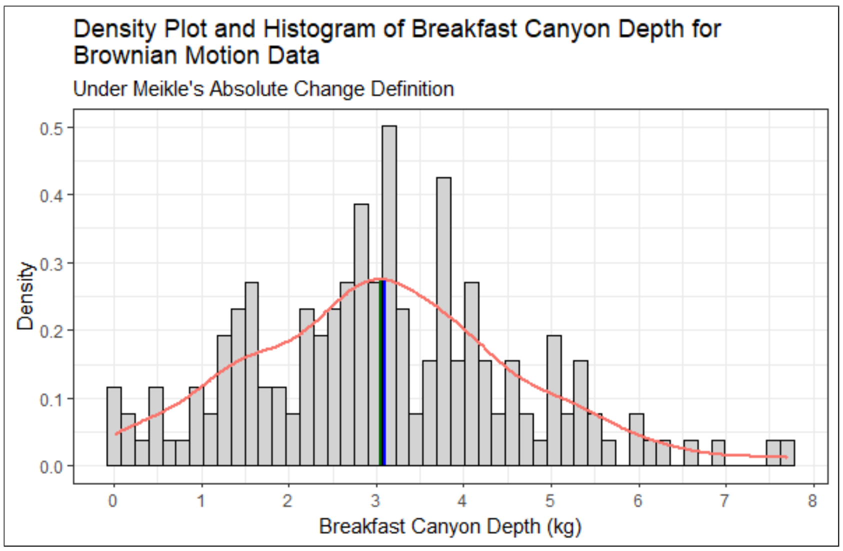

Figure 3.14: Pictured is a histogram of the distribution of the Breakfast Canyon Depth for the simulated Brownian Motion data, using Meikle's absolute slope definition for Breakfast Canyon observation, with the density curve plotted in red. The median of the Breakfast Canyon Depth is the green line, while the mean of the Breakfast Canyon Depth is the blue line.

The onset of the Breakfast Canyon begins at the 1st breakpoint 38 of the hive-day instances. All of the 38 hive-day isntances were instances where a 2 breakpoint SLR model was chosen. The onset of the Breakfast Canyon began at the 2nd breakpoint 127 of the hive-day instances, and all of which were only instances in which a 3 breakpoint SLR model was chosen. A summary of the location of the onset of the Breakfast Canyon is provided in Table 3.6.

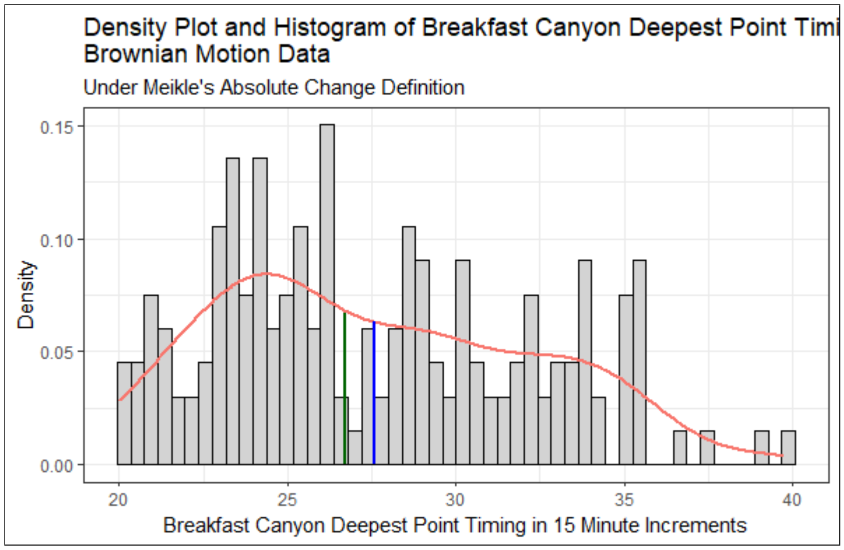

Figure 3.15: Pictured is a histogram of the distribution of the time of the deepest point of the Breakfast Canyon for the simulated Brownian Motion data, using Meikle's absolute slope definition for Breakfast Canyon observation, with the density curve plotted in red. The median of the time of the deepest point is the green line, while the mean of the time of the deepest point is the blue line.

The mean time for Breakfast Canyon onset for the Brownian Motion data under Meikle's absolute slope framework is around 4:15 AM, with a standard deviation of around 58 minutes, and a median Breakfast Canyon onset of around 4:01 AM. A visualization of the distribution of the Breakfast Canyon Onset timing is shown in Figure 3.16.

Lastly, we plot the extracted line segments for the observed Breakfast Canyon Brownian Motion instances from the onset of the Breakfast Canyon to the deepest point of the canyon. We normalized the data to have all segments begin at zero, allowing us to visualize the depth of each Breakfast Canyon Brownian Motion instance, by Pattern type, visualized in Figure 3.17

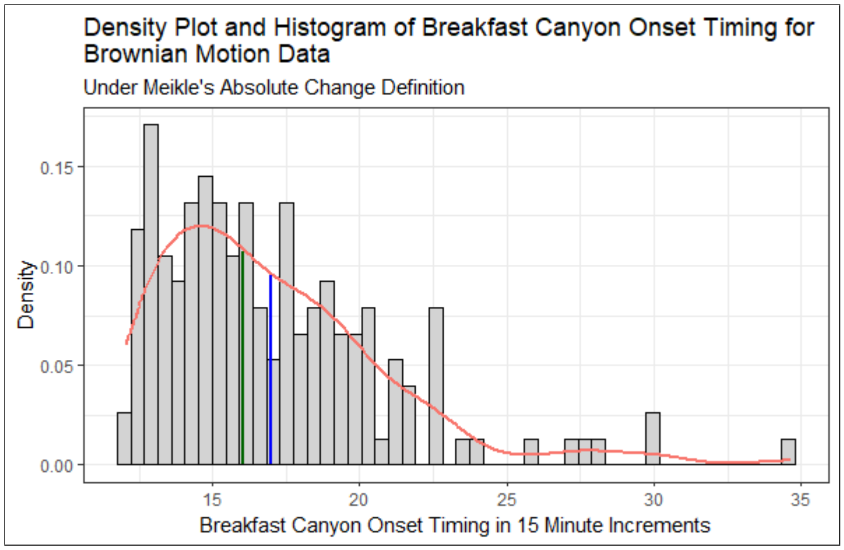

Figure 3.16: Pictured is a histogram of the distribution of the time of the onset of the Breakfast Canyon for the simulated Brownian Motion data, using Meikle's absolute slope definition for Breakfast Canyon observation, with the density curve plotted in red. The median of the time of the Breakfast Canyon onset is the green line, while the mean of the time of the Breakfast Canyon onset is the blue line.

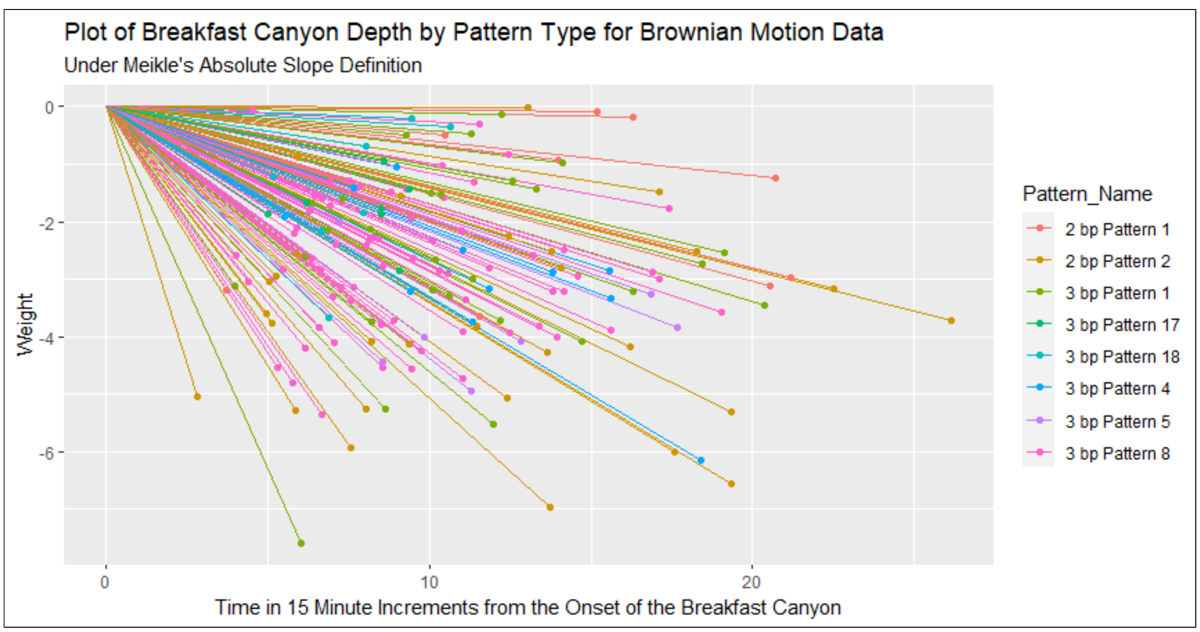

Figure 3.17: Pictured are the extracted line segments used to estimate the Breakfast Canyon depth in kg for Brownian Motion instances with an observed Breakfast Canyon (n=165) under Meikle's Absolute slope definition. The segments were normalized to begin at zero weight, with time at zero being the time of the onset of the Breakfast Canyon. Line segments and endpoints were color-coded by Pattern Type.

3.2.2.2 Results for Brownian Motion Under Our Relative Slope Breakfast Canyon Definition

Under our novel relative slope Breakfast Canyon definition, 202 of the 1,000 hive-day instances had an observed Breakfast Canyon. Thus, 20.2% of the hive-day instances had posterior mean slope patterns which indicated a Breakfast Canyon was present. Of these 202 hive-day instances, 46 of the hive-day had a chosen 2 breakpoint Segmented Linear Regression model, while the remaining 156 had a chosen 3 breakpoint Segmented Linear Regression Model.

The 202 hive-day instances where a Breakfast Canyon was observed fall into 23 of the 41 possible Breakfast Canyon patterns between the 2 breakpoint and 3 breakpoint relative slope patterns that result in an observed Breakfast Canyon in the data. A numerical summary of the Breakfast Canyon patterns observed under a relative slope framework are shown in Table 3.7. In the 2 breakpoint SLR model framework, only one Breakfast Canyon pattern was not observed in the simulated Brownian motion data, which was Pattern 11. In the 3 breakpoint SLR model framework, 17 Breakfast Canyon patterns were not observed in the simulated Brownian motion data, which include Patterns 1, 2, 3, 4, 5, 6, 10, 12, 21, 28, 33, 36, 41, 45, 46, 51 and 53. Thus, a total of 18 relative slope patterns with an observed Breakfast canyon were not present in the simulated Brownian motion data. The most common Breakfast Canyon pattern observed in the Brownian Motion data was pattern 17 for a 3 breakpoint SLR model. After Pattern 17, patterns 15, 39, 30 and 40 from a 3 breakpoint SLR model, and Pattern 2 from a 2 breakpoint SLR model were most commonly observed in the Brownian Motion data. Posterior mean plots for one hive-day instance following each pattern for the top 5 observed Breakfast Canyon patterns can be shown in Figure 3.18.

The Breakfast Canyon reached its' depth at the 2nd breakpoint 46 of the hive-day instances. All of the 46 hive-day instances were instances where a 2 breakpoint SLR model was chosen. The Breakfast Canyon reached its' depth at the 3rd breakpoint 156 of the hive-day instances, all of which could only be instances in which a 3 breakpoint SLR model was chosen. A summary of the breakpoint location of the canyon depth for days with an observed Breakfast Canyon can be found in Table 3.8.

The mean Breakfast Canyon depth for the Brownian Motion data under our relative slope definition framework is 3.060 kg, with a standard deviation of 1.598 kg. The median

Model Selected	**Pattern**	**Number of Instances**
3 breakpoint	Pattern 17	67
2 breakpoint	Pattern 2	29
3 breakpoint	Pattern 15	16
3 breakpoint	Pattern 39	15
3 breakpoint	Pattern 40	13
3 breakpoint	Pattern 30	11
2 breakpoint	Pattern 7	7
3 breakpoint	Pattern 18	7
2 breakpoint	Pattern 10	6
3 breakpoint	Pattern 20	5
3 breakpoint	Pattern 7	5
2 breakpoint	Pattern 9	3
3 breakpoint	Pattern 31	3
3 breakpoint	Pattern 38	3
3 breakpoint	Pattern 19	2
3 breakpoint	Pattern 25	2
3 breakpoint	Pattern 29	2
2 breakpoint	Pattern 17	1
3 breakpoint	Pattern 11	1
3 breakpoint	Pattern 14	1
3 breakpoint	Pattern 34	1
3 breakpoint	Pattern 48	1
3 breakpoint	Pattern 54	1

Table 3.7: Described in the table above are the number of simulated Brownian Motion instances that fall into the 23 observed Breakfast Canyon patterns in our simulated Brownian Motion data under our relative slope Breakfast Canyon definition.

Model Selected	**Location of the Bottom of the Breakfast Canyon**		**Location of the Onset of the Breakfast Canyon**	
	2nd Breakpoint	*3rd Breakpoint*	*1st Breakpoint*	*2nd Breakpoint*
2 breakpoint SLR model	46		46	
3 breakpoint SLR model	0	156	0	156

Table 3.8: A summary of the counts of the locations of the bottom of the Breakfast Canyon for Brownian Motion simulated data under our novel relative slope Breakfast Canyon definition, as well as the locations of the onset of the Breakfast Canyon. Counts are broken down by the model that was selected via WAIC, which is either the 2 or 3 breakpoint Segmented Linear Regression Model.

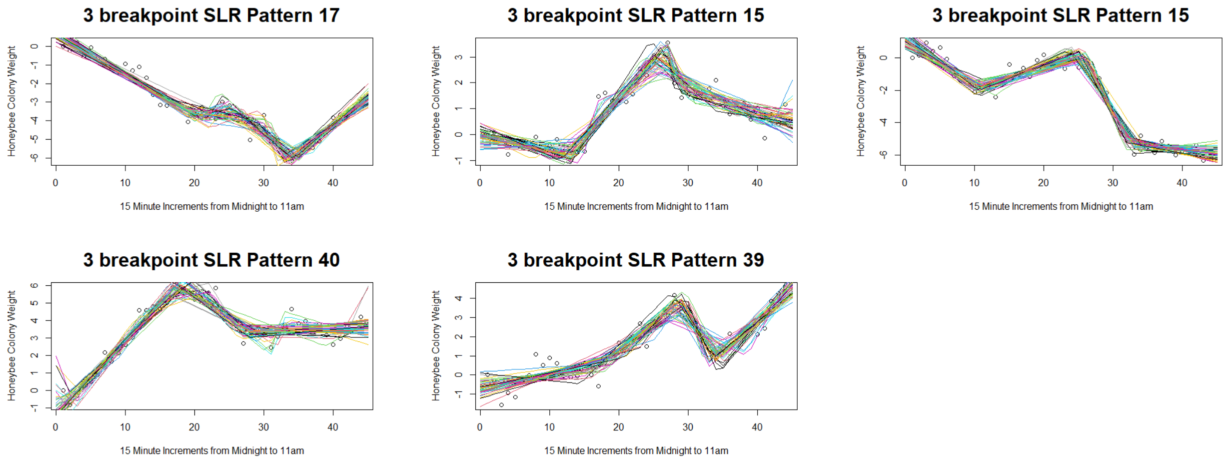

Figure 3.18: For the top 5 patterns with an observed Breakfast Canyon using our relative slope Breakfast Canyon definition, we chose one simulated brownian motion instance per pattern. We plot the raw simulated brownian motion instance, shown as the points on the plot. Further, we plot 100 means from the posterior distribution, which are the colored lines of the plot.

Breakfast Canyon depth for the Brownian Motion data is 3.023 kg, and the distribution of the Breakfast Canyon Depth is roughly symmetric, shown in Figure 3.19.

The mean time for Breakfast Canyon's deepest point for the Brownian Motion data is around 6:49 AM, with a standard deviation of 1 hour 7 minutes, and a median timing of the Breakfast Canyon's deepest point of around 6:35 AM. A visualization of the distribution of the Breakfast Canyon deepest point timing is shown in Figure 3.20.

The onset of the Breakfast Canyon begins at the 1st breakpoint 46 of the hive-day instances.All of the 46 hive-day instances were instances where a 2 breakpoint SLR model was chosen. The onset of the Breakfast Canyon began at the 2nd breakpoint 156 of the hive-day instances, and all of which were only instances in which a 3 breakpoint SLR model was chosen. A summary of the location of the onset of the Breakfast Canyon is provided in Table 3.8. 15 of the day-hive instances with the onset beginning at the first breakpoint reached the bottom of the canyon at the 3rd breakpoint, meaning that the Breakfast Canyon was comprised of the 2nd, 3rd, and 4th segments.

The mean time for Breakfast Canyon onset for the Brownian Motion data is around 4:14 AM, with a standard deviation of 59 minutes, and a median Breakfast Canyon onset

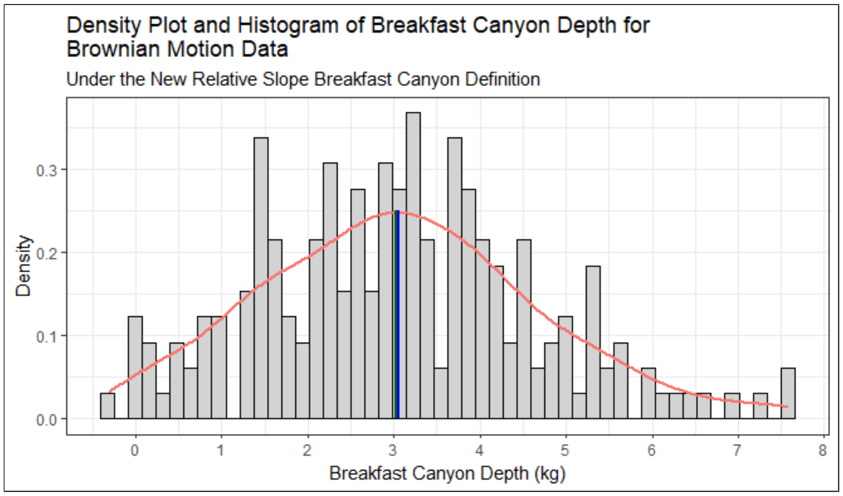

Figure 3.19: Pictured is a histogram of the distribution of the Breakfast Canyon Depth for the Real simulated Brownian Motion data, using our relative slope definition for Breakfast Canyon observation, with the density curve plotted in red. The median of the Breakfast Canyon Depth is the green line, while the mean of the Breakfast Canyon Depth is the blue line.

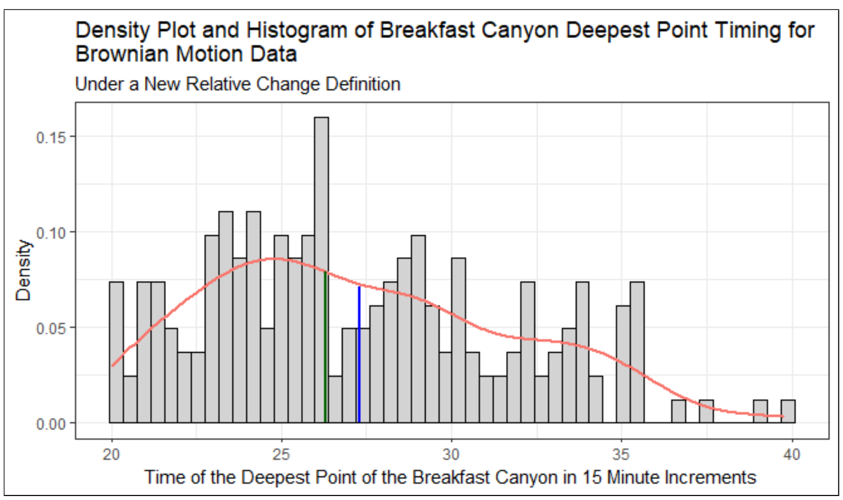

Figure 3.20: Pictured is a histogram of the distribution of the time of the deepest point of the Breakfast Canyon for the simulated Brownian Motion data, using Our relative slope definition for Breakfast Canyon observation, with the density curve plotted in red. The median of the time of the deepest point is the green line, while the mean of the time of the deepest point is the blue line.

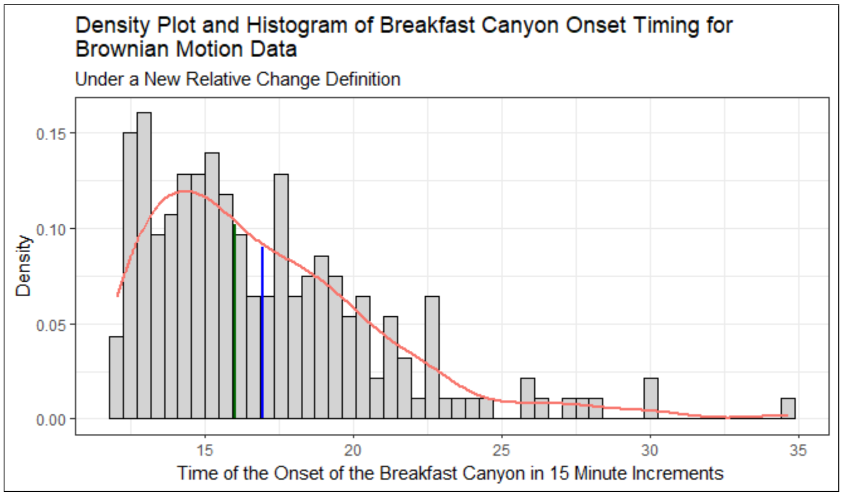

Figure 3.21: Pictured is a histogram of the distribution of the time of the onset of the Breakfast Canyon for the simulated Brownian Motion data, using our relative slope definition for Breakfast Canyon observation, with the density curve plotted in red. The median of the time of the Breakfast Canyon onset is the green line, while the mean of the time of the Breakfast Canyon onset is the blue line.

of around 4:00 AM. A visualization of the distribution of the Breakfast Canyon Onset timing is shown in Figure 3.21.

Lastly, we plot the extracted line segments for the observed Breakfast Canyon Brownian Motion instances from the onset of the Breakfast Canyon to the deepest point of the canyon. We normalized the data to have all segments begin at zero, allowing us to visualize the depth of each Breakfast Canyon Brownian Motion instance, by Pattern type, visualized in Figure 3.22

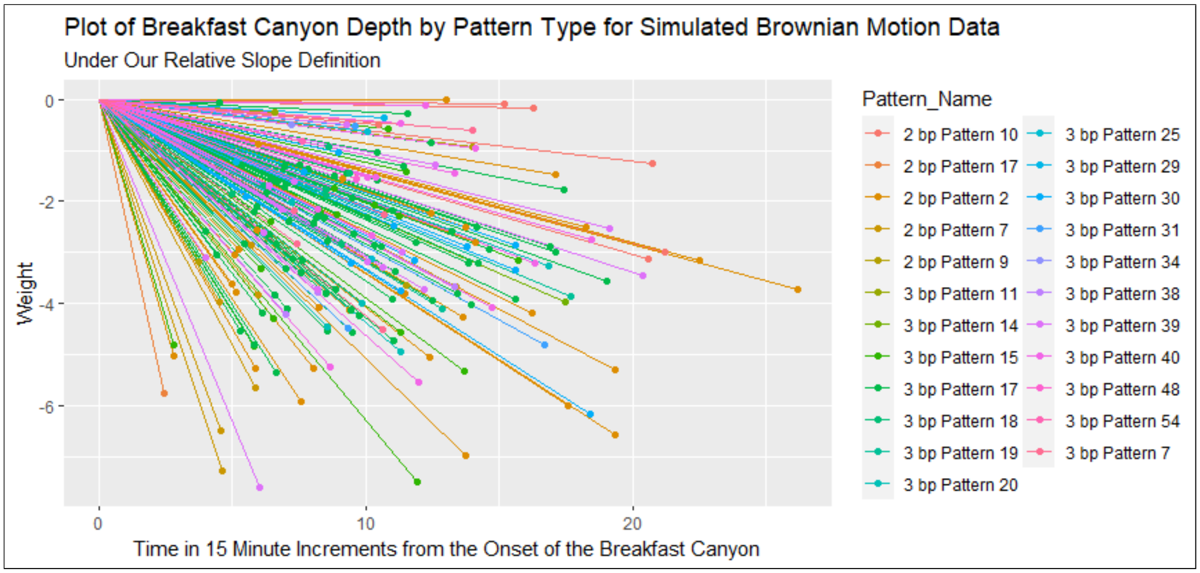

Figure 3.22: Pictured are the extracted line segments used to estimate the Breakfast Canyon depth in kg for Brownian Motion instances with an observed Breakfast Canyon (n=202) under Our relative slope definition. The segments were normalized to begin at zero weight, with time at zero being the time of the onset of the Breakfast Canyon. Line segments and endpoints were color-coded by Pattern Type.

3.2.2.3 A Comparison of the Results Between Breakfast Canyon Definitions for Simulated Brownian Motion Data

We next compare the results of applying Meikle's absolute slope definition and Our new relative slope definition to simulated Brownian Motion data. We first observe that Meikle's absolute slope definition yields a proportion of 16.5% of Brownian Motion instances with an observed Breakfast Canyon, while Our relative slope definition yields a proportion of 20.2% of Brownian Motion instances with an observed Breakfast Canyon. Since our patterns extend upon Meikle's patterns, a larger proportion of observed Breakfast Canyon Brownian Motion instances is what we would expect to see theoretically.

We observe when comparing the distributions of the Breakfast Canyon Depth in Kilograms between the two definitions (Figures 3.19 and 3.14) that they both result in symmetric distributions. Similarly, the distributions of the timing of the deepest point of the Breakfast Canyon are visually similar when comparing between the two definitions (Figures 3.20 and 3.15). Further, the distributions of the timing of the onset of the Breakfast Canyon are visually similar when comparing the two definitions (Figures 3.21 and 3.16). Further, when comparing the extracted segments that are used to estimate

Breakfast Canyon depth, shown in Figures 3.17 and 3.22, some of the extracted line segments have a shown positive slope under Our relative definition, while none of the extracted line segments have a shown negative slope under our relative definition, due to the extension of our definition.

3.2.3 A Comparison Between Our Methods on Real Honeybee Data and Simulated Brownian Motion Data

We applied our methods to both Real Honeybee data and simulated Brownian Motion data to explore the existence of the Breakfast Canyon in our data. We wish to establish a framework that shows that real Honeybee data results in Breakfast Canyon patterns in a greater proportion than random noise does, which is why we compare our results against simulated Brownian Motion data. Additional research should be conducted, and statistical tests should be established beyond what is described in this section, as this is just the first step.

In the Honeybee data, we found that under Meikle's absolute slope definition, 5.27% of the hive-day instances had an observed Breakfast Canyon. Under Our relative slope definition, 8.83% of the hive-day instances had an observed Breakfast Canyon. In the Brownian motion data, we found that under Meikle's absolute slope definition, 16.5% of the instances had an observed Breakfast Canyon. Under Our relative slope definition, 20.2% of the instances had an observed Breakfast Canyon. From our results, we find that the simulated Brownian Motion data resulted in a higher proportion of observed Breakfast Canyons under both definitions. The real Honeybee data resulted in a lower proportion of observed Breakfast Canyons under both definitions, yielding the question of whether a Breakfast Canyon pattern truly exists in the hive weight data. Further exploration with more data sets would be needed to answer this question fully, as other factors such as equipment error could impact the weight scale data.

Of the Breakfast Canyon hive-day instances observed under Meikle's definition using the Honeybee data, 17.39% of the hive-day instances resulted in a 2 breakpoint SLR model selected by WAIC, while 82.61% of the hive-day instances resulted in a 3 breakpoint SLR model selected by WAIC. Under our absolute slope definition using the Honeybee data, 13.55% of the instances had a selected 2 breakpoint SLR model, and 86.45% of the instances had a selected 3 breakpoint SLR model. Of the Breakfast Canyon instances resulting in a Breakfast Canyon pattern under Meikle's definition using simulated Brownian Motion data, 23.03% of the instances had a selected 2 breakpoint SLR model, while 76.97% of the instances had a selected 3 breakpoint SLR model. Under our absolute slope definition using the simulated Brownian motion data, 22.77% of the instances had a selected 2 breakpoint SLR model, and 77.23% of the instances had a selected 3 breakpoint SLR model. The proportion of instances in both the Brownian motion data and the Honeybee data that respectively chose a 2 breakpoint SLR model or a 3 breakpoint SLR model are similar.

Lastly, we discuss our observations when comparing the Breakfast Canyon Depth results between the real Honeybee data and the simulated Brownian Motion data. Since the real Honeybee data and simulated Brownian Motion are not on the same scale, we will not compare the measures of center such as the mean and median. Instead, we observe the shape of the distribution of the Breakfast Canyon depth (Figures 3.7, 3.2, 3.19 and 3.14). As discussed in sections 3.1.2.3 and 3.2.2.3, between definitions the distribution of the Breakfast Canyon depths are similar. However, when comparing the real Honeybee data Breakfast Canyon depth distribution to the simulated Brownian Motion data Breakfast Canyon depth distribution, we observe differences in the shape of the distribution. The distribution of the Breakfast Canyon depth is right skewed for both definitions in the Honeybee data, while the distribution of the Breakfast Canyon depth is roughly symmetric for both definitions in the simulated Brownian Motion data.

Chapter 4
Discussion and Future Work

In this master's thesis, we have used a segmented linear regression model to estimate the linear trends and breakpoints of Honeybee hive weight data. We wished to expand upon Holst and Meikle's work, and begin exploring the existence of a Breakfast Canyon in our own data (2018). We created a Bayesian segmented linear regression model with 3 breakpoints, and a Bayesian segmented linear regression model with 2 breakpoints. We introduce a novel Breakfast Canyon definition, that uses a relative slope framework instead of an absolute slope framework. Furthermore, we suggest an information criteria such as WAIC to more effectively compare models with varying numbers of breakpoints instead of R-squared, which was used for model selection in Meikle and Holst's 2018 paper. WAIC will help to identify the effective number of breakpoints to adjust for overfitting.

While we were successful in creating a fully Bayesian SLR for both 2 and 3 breakpoints, as well as a novel Breakfast Canyon definition, further exploration and validation of this method is needed. We based the slopes corresponding to both Meikle's absolute slope and our relative slope Breakfast Canyon definitions based on the posterior means of the slope parameters, $\hat{\alpha_1}, \hat{\alpha_2}, \hat{\alpha_3}$ and $\hat{\alpha_4}$ for the 3 breakpoint SLR model, and $\hat{\alpha_1}, \hat{\alpha_2}$ and $\hat{\alpha_3}$ for the 2 breakpoint SLR model. After completing our analysis, we chose to review the 95% credible intervals of the slope parameters for a few days for both the real Honeybee data and the Brownian Motion data. If the 95% credible intervals for the slope parameters contain 0, it means that the slope in question could potentially be negative or positive. We first will discuss our findings when analyzing the real Honeybee data. We chose 12 hive-day instances, some of which a 2 breakpoint SLR model was the better fit, and some of which a 3 breakpoint SLR model was the better fit. We found the 95% credible interval of the slope parameters for those 12 days, pictured in Figure 4.1.

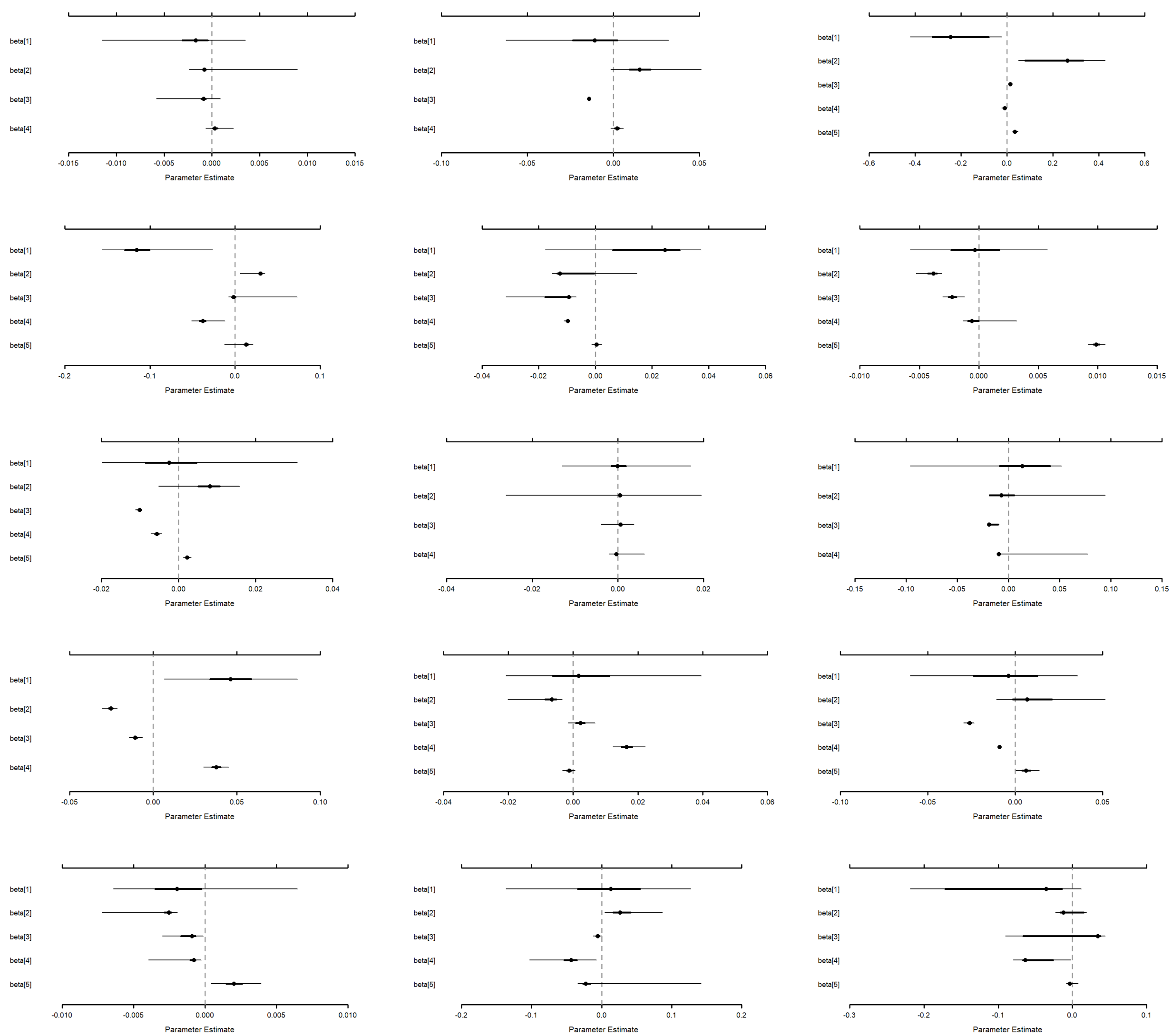

Figure 4.1: 95% Credible Intervals for the regression parameters $\beta = (\mu_0, \alpha_1, \alpha_2, \alpha_3, \alpha_4)'$ if the model chosen was a 3 breakpoint SLR model, $\beta = (\mu_0, \alpha_1, \alpha_2, \alpha_3)'$ if the model chosen was a 2 breakpoint SLR model. The point represents the posterior median, the thick line is the 50% credible interval, and the thin line is the 95% credible interval for the slope parameter. The data in question is from real Honeybee hive weight data.

The intercept parameter, μ_0 is the only parameter where the 95% credible interval could contain zero. When we do not take into account the intercept parameters, there are 55 slope parameters between the 12 days, and we see that 24 of the slope parameters have 95% credible intervals that contain zero.

We next discuss our findings when analyzing the data simulated from Brownian motion. We chose 11 simulated Brownian motion instances, some of which where a 2

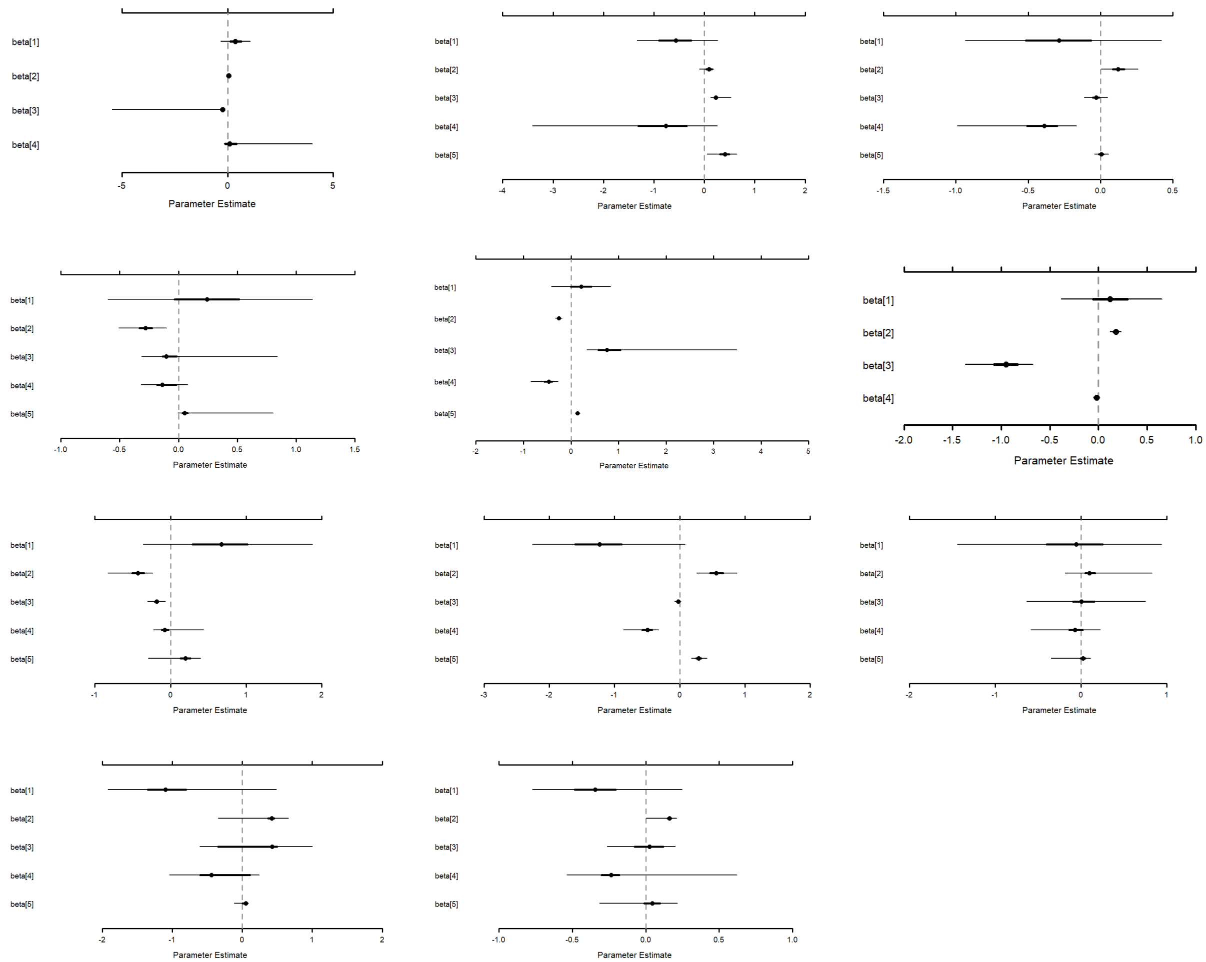

Figure 4.2: 95% Credible Intervals for the regression parameters $\beta = (\mu_0, \alpha_1, \alpha_2, \alpha_3, \alpha_4)'$ if the model chosen was a 3 breakpoint SLR model, $\beta = (\mu_0, \alpha_1, \alpha_2, \alpha_3)'$ if the model chosen was a 2 breakpoint SLR model. The point represents the posterior median, the thick line is the 50% credible interval, and the thin line is the 95% credible interval for the slope parameter. The data in question is from simulated Brownian Motion.

breakpoint SLR model was the better fit, and some of which where a 3 breakpoint SLR model was the better fit. We found the 95% credible interval of the slope parameters for those 11 instances, pictured in Figure 4.2. The intercept parameter, μ_0 is the only parameter where the 95% credible interval could contain zero. When we do not take into account the intercept parameters, there are 42 slope parameters between the 11 instances, and we see that 23 of the slope parameters have 95% credible intervals that contain zero.

For both the real Honeybee data and the simulated Brownian Motion data, over half of the slope parameters include zero in their 95% credible intervals, which poses some issues in terms of the validity of this method, as this means that the slopes could be positive or negative. More research is needed to address this issue, potentially with using each iteration to compute the posterior probability of the observation of a Breakfast Canyon in the data instead of utilizing the posterior means as the basis of our method to observe Breakfast Canyon in data.

When comparing our results utilizing the methods established on Brownian Motion data and real Honeybee data from Michigan, we found Brownian Motion resulted in a higher proportion of observed Breakfast Canyons. This suggests that the real Honeybee data that we analyzed could have been classified as having Breakfast Canyons simply due to random noise. In future work, a more robust method to compare real Honeybee data to random noise could be established to provide further support that the Breakfast Canyon exists. Replication of our analysis with different real Honeybee weight data is needed to further explore the existence of the Breakfast Canyon.

One of our main objectives in this paper was to focus on creating a more concrete definition of a Breakfast Canyon. In future work, after running our Bayesian SLR model on all of the days and finding if there was a Breakfast Canyon observed or not, we could include factors such as precipitation, temperature, and humidity to explore if these factors impact the presence of a Breakfast Canyon.

In this thesis, we established methods to aid apiarists and researchers in exploring the existence of Breakfast Canyon patterns and the subsequent Breakfast Canyon depth in Honeybee hive weight data, which can give insights into the foraging force of Honeybee colonies. We established a 2 and 3 breakpoint Bayesian Segmented Linear Regression model, utilizing WAIC as our criteria for model selection. We establish a framework for observing Breakfast Canyon patterns in hive weight data based on Meikle's absolute slope Breakfast Canyon definition, and create a novel framework based on a relative slope Breakfast Canyon definition. We estimate the Breakfast Canyon depth for days in which a Breakfast Canyon pattern is observed, contributing the the understanding of foraging force, and subsequently Honeybee colony health and performance.

Appendix
Figures, Tables and R Code

```
## Load in data

lgood_days2 <- readRDS("good_foraging_days.rdata")
num_unique=4070
lgood_days <- vector(mode='list', length=num_unique)
for(i in 1:4070){
  one_day <- lgood_days2[[i]]
  one_day2 <- filter(one_day, hour <=11)
  lgood_days[[i]] <- one_day2
}


lBC_Analysis <- vector(mode='list', length=num_unique)


## Extra Functions
choose_mean <- nimbleFunction( run = function(beta1=double(),
beta2=double(),beta3=double(),beta4=double(),beta5=double(),
bp1=double(),bp2=double(), bp3=double(), obs=double()){
  if(obs <= bp1){ mean <- beta1+beta2*obs}

  else if(obs > bp1 & obs <= bp2){mean <- beta1+beta2*bp1
  +beta3*(obs-bp1)}
```

```
  else if(obs > bp2 & obs <= bp3){mean <- beta1+beta2*bp1
  +beta3*(bp2-bp1)+beta4*(obs-bp2)}

  else if(obs > bp3){
    mean <- beta1+beta2*bp1+beta3*(bp2-bp1)
    +beta4*(bp3-bp2)+beta5*(obs-bp3)}

  return(mean)
  returnType(double())
})


get.mean=function(beta,bp, obs){
  beta1=beta[1]
  beta2=beta[2]
  beta3=beta[3]
  beta4=beta[4]
  beta5=beta[5]
  bp1=bp[1]
  bp2=bp[2]
  bp3=bp[3]
  if(obs <= bp1){ mean <- beta1+beta2*obs}

  else if(obs > bp1 & obs <= bp2){mean <- beta1+beta2*bp1+beta3*(obs-bp1)

  else if(obs > bp2 & obs <= bp3){mean <- beta1+beta2*bp1
  +beta3*(bp2-bp1)+beta4*(obs-bp2)}

  else if(obs > bp3 ){
    mean <- beta1+beta2*bp1+beta3*(bp2-bp1)
    +beta4*(bp3-bp2)+beta5*(obs-bp3)}

  return(mean)

}
```

```
## Big Function
for(k in 3910:num_unique){
  one_day <- lgood_days[[k]]
  Y=one_day$new_weight
  ## 3 breakpoints
  M=5 ## number of regression parameters
  T=length(Y) ## Set T
  bps = c(10,18,25) ## bp initials

  ## Nimble Code
  model <- nimbleCode({
    beta[1]~ dnorm(mean=0,sd=100)
    ## note this is technically mu_0
    for(a in 2:M){
      beta[a] ~ dnorm(mean = 0, sd=100)
      ## note these are alpha_{1,2,3,4}
    }

    sigma2 ~ dexp(1)

    for(t in 1:T){
      mean[t] <-  choose_mean(beta1=beta[1],beta2=beta[2],beta3=beta[3],
      beta4=beta[4],beta5=beta[5], bp1=bp[1],
      bp2=bp[2], bp3=bp[3], obs=t)
    }

    for(t in 1:T){
      Y[t] ~dnorm(mean=mean[t],sd=sqrt(sigma2))
    }

    omega[1:4]~ddirch(alphas[1:4])

    bp[1]<-(omega[1]*T)
```

```
  bp[2]<-(omega[1]+omega[2])*T
  bp[3]<-(omega[1]+omega[2]+omega[3])*T
})
## Model Data
model.data <- list(
  Y = Y
)
## Setting Constants
model.constants <- list(
  T=T,
  M=M, ## number of betas 1...5
  alphas=c(1,1,1,1)*3
)
## Setting Initial Values
model.inits <- list(
  sigma2 = 0.01,
  beta = rep(0, M),
  omega=c(.2,.2,.4,.2),
  bp=bps
)
## Run the MCMC
mod <- nimbleModel(
  model,
  data = model.data,
  constants = model.constants,
  inits = model.inits
)
mod$initializeInfo()
## Compile
comp.model <- compileNimble(
  mod
)
## Configure
conf.model <- configureMCMC(
  comp.model,
```

```
  monitors = c("sigma2","beta","bp","omega"),
  enableWAIC = TRUE
)
## Build
build.mcmc <- buildMCMC(
  conf.model
)
## Compile
c.mcmc <- compileNimble(
  build.mcmc
)
## Run MCMC
model.out <- runMCMC(
  c.mcmc,
  niter = 100000,
  nburnin = 1000
)
## Calculate WAIC
WAIC <- nimble::calculateWAIC(
  mcmc = model.out,
  model = mod
)$WAIC
print(WAIC)


## Get posterior means 3 bp
post.means=apply(model.out,2,mean)
post.means

beta.hat=post.means[1:5] ## Posterior means for betas
bp.hat=post.means[6:8]## Posterior means for break points

## Meikle's definition of breakfast canyon satisfied?
## Do the bps exist at the proper time?
## bp3 between 5-10 (20-40 adjusted)
```

```
## Detect if breakpoint timing is right
time_right_3 <- c()
time_right_2 <- c()

 if(bp.hat[3] > 20 & bp.hat[3] <40){time_right_3 = 1}
 else{ time_right_3=0}
 if(bp.hat[2] > 20 & bp.hat[2] <40){time_right_2 = 1}
 else{ time_right_2=0}

## Get yhat
bp1y.hat= get.mean(beta=beta.hat,bp=bp.hat,obs=bp.hat[1])
bp2y.hat= get.mean(beta.hat,bp.hat,bp.hat[2])
bp3y.hat= get.mean(beta.hat,bp.hat,bp.hat[3])
T.hat= get.mean(beta.hat,bp.hat,44)

BC_Analysis <- data.frame(bp1y.hat,bp2y.hat,bp3y.hat)
BC_Analysis$BC_Depth_3 <- BC_Analysis$bp2y.hat - BC_Analysis$bp3y.hat
BC_Analysis$BC_Depth_2 <- BC_Analysis$bp1y.hat - BC_Analysis$bp2y.hat
BC_Analysis$In_Time_3 <- time_right_3
BC_Analysis$In_Time_2 <- time_right_2
BC_Analysis$WAIC <- WAIC
BC_Analysis$T.hat <- T.hat

 regain_3 <- c()
 regain_2 <- c()
 if(BC_Analysis$T.hat > BC_Analysis$bp3y.hat){regain_3 = 1}
 else {regain_3 = 0}
 if(BC_Analysis$bp3y.hat > BC_Analysis$bp2y.hat){regain_2 = 1}
 else {regain_2 = 0}

BC_Analysis$regain_3 <- regain_3
BC_Analysis$regain_2 <- regain_2

 loss_3 <- c()
```

```
  loss_2 <- c()

  if(BC_Analysis$bp3y.hat>BC_Analysis$bp2y.hat){loss_3=1} else{loss_3=0}
  if(BC_Analysis$bp2y.hat>BC_Analysis$bp1y.hat){loss_2=1} else{loss_2=0}

  BC_Analysis$loss_3 <- loss_3
  BC_Analysis$loss_2 <- loss_2



  BC_Exists_3 <- c()
  BC_Exists_2 <- c()

if(BC_Analysis$In_Time_3 == "1" &
BC_Analysis$regain_3 == "1" & BC_Analysis$loss_3 == "1")
{BC_Exists_3=1} else{BC_Exists_3=0}

if(BC_Analysis$In_Time_2 == "1" &
BC_Analysis$regain_2 == "1" & BC_Analysis$loss_2 == "1")
{BC_Exists_2=1} else{BC_Exists_2=0}

  BC_Analysis$BC_Exists_2 <- BC_Exists_2
  BC_Analysis$BC_Exists_3 <- BC_Exists_3
  BC_Analysis$unique <- rep(one_day$unique[1],51)

  BC_Detected <- c()
if(BC_Analysis$BC_Exists_3=="1" | BC_Analysis$BC_Exists_2=="1")
{BC_Detected=1} else {BC_Detected=0}
  BC_Analysis$BC_Detected <- BC_Detected
  unique <- unique(one_day$totally_unique)
  BC_Analysis$totally_unique <- unique
  BC_Analysis$Omega1 <- post.means[7]
  BC_Analysis$Omega2 <- post.means[8]
  BC_Analysis$Omega3 <- post.means[9]
  BC_Analysis$Omega4 <- post.means[10]
  BC_Analysis$bp1 <- post.means[5]
```

```
  BC_Analysis$bp2 <- post.means[6]
  BC_Analysis$bp3 <- post.means[7]
  BC_Analysis$beta1 <- post.means[1]
  BC_Analysis$beta2 <- post.means[2]
  BC_Analysis$beta3 <- post.means[3]
  BC_Analysis$beta4 <- post.means[4]
  BC_Analysis$beta5 <- post.means[5]


  output <- BC_Analysis[c(1:3,4,5,8,9,14,15,16,17:29)]
  lBC_Analysis[[k]] <- output

  }

###################################
## End for loop
BC_numerator <- c()
for(i in 1:num_unique){
  day <- lBC_Analysis[[i]]
  value = day$BC_Detected
  if(value == 1){BC_numerator[i] <- 1} else {BC_numerator[i] <- 0}
}

numerator_3 <- sum(sums_3 >=26, na.rm=TRUE)
numerator_2 <- sum(sums_2 >=26, na.rm=TRUE)
numerator = numerator_2 + numerator_3
denominator <- 100
prop_BC <- numerator/denominator
```

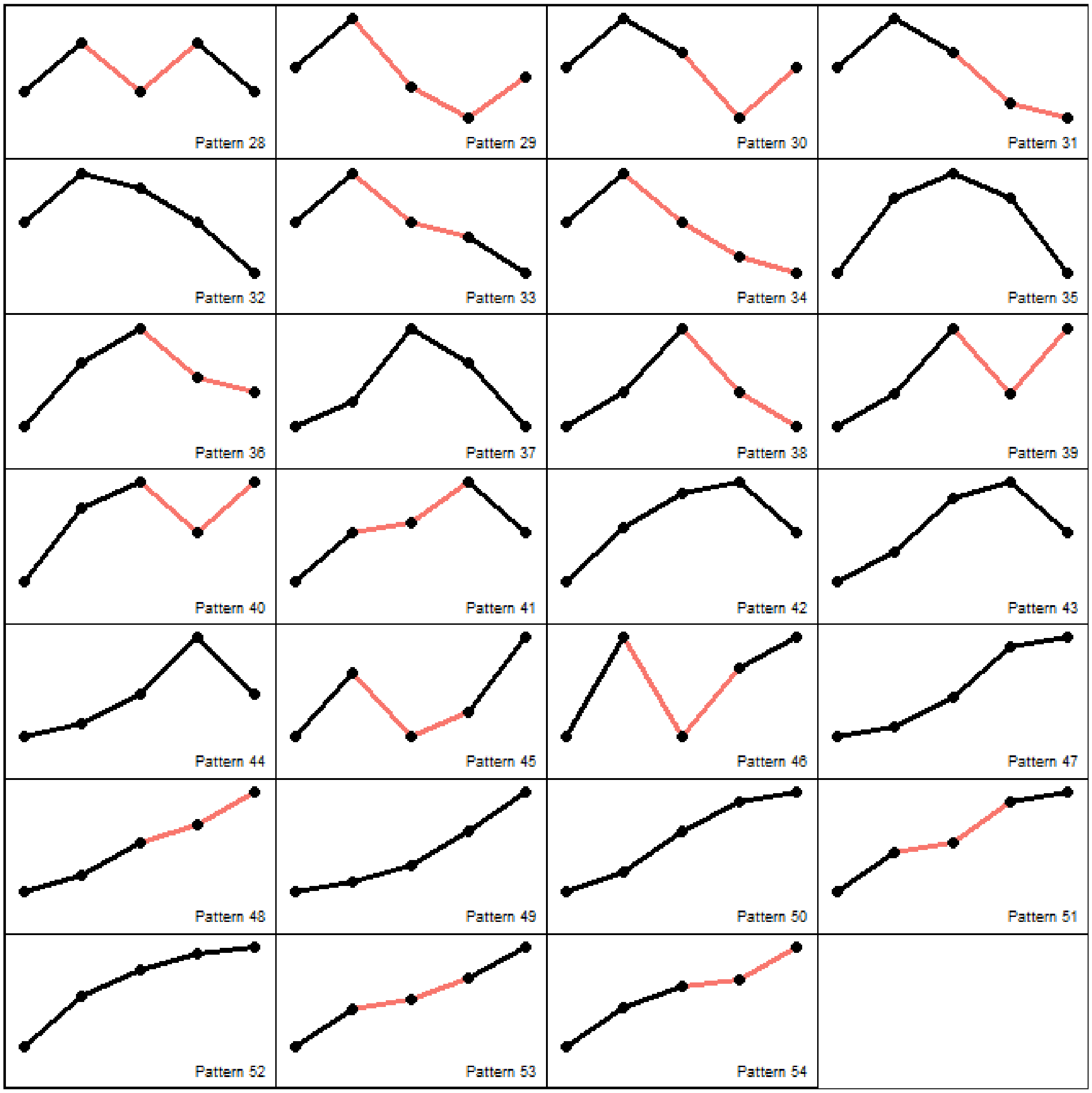

Figure .1: This figure gives a visualization that corresponds with table .1 of the second 27 of 54 possible relative change slope patterns using our 3 breakpoint Segmented Linear Regression model Breakfast Canyon definition.

Pattern	**Segment 1**	**Segment 2**	**Segment 3**	**Segment 4**
Pattern 28	$\hat{\alpha_1} > 0$	$\hat{\alpha_2} < 0$	$\hat{\alpha_3} > 0$	$\hat{\alpha_4} < 0$
Pattern 29	$\hat{\alpha_1} > 0$	$\hat{\alpha_2} < 0$	$\hat{\alpha_3} < 0$ and $\hat{\alpha_3} > \hat{\alpha_2}$	$\hat{\alpha_4} > 0$
Pattern 30	$\hat{\alpha_1} > 0$	$\hat{\alpha_2} < 0$	$\hat{\alpha_3} < 0$ and $\hat{\alpha_3} \leq \hat{\alpha_2}$	$\hat{\alpha_4} > 0$
Pattern 31	$\hat{\alpha_1} > 0$	$\hat{\alpha_2} < 0$	$\hat{\alpha_3} < 0$ and $\hat{\alpha_3} < \hat{\alpha_2}$	$\hat{\alpha_4} < 0$ and $\hat{\alpha_4} > \hat{\alpha_3}$
Pattern 32	$\hat{\alpha_1} > 0$	$\hat{\alpha_2} < 0$	$\hat{\alpha_3} < 0$ and $\hat{\alpha_3} < \hat{\alpha_2}$	$\hat{\alpha_4} < 0$ and $\hat{\alpha_4} \leq \hat{\alpha_3}$
Pattern 33	$\hat{\alpha_1} > 0$	$\hat{\alpha_2} < 0$	$\hat{\alpha_3} < 0$ and $\hat{\alpha_3} > \hat{\alpha_2}$	$\hat{\alpha_4} < 0$ and $\hat{\alpha_4} \leq \hat{\alpha_3}$
Pattern 34	$\hat{\alpha_1} > 0$	$\hat{\alpha_2} < 0$	$\hat{\alpha_3} < 0$ and $\hat{\alpha_3} > \hat{\alpha_2}$	$\hat{\alpha_4} < 0$ and $\hat{\alpha_4} > \hat{\alpha_3}$
Pattern 35	$\hat{\alpha_1} > 0$	$\hat{\alpha_2} > 0$ and $\hat{\alpha_2} < \hat{\alpha_1}$	$\hat{\alpha_3} < 0$	$\hat{\alpha_4} < 0$ and $\hat{\alpha_4} \leq \hat{\alpha_3}$
Pattern 36	$\hat{\alpha_1} > 0$	$\hat{\alpha_2} > 0$ and $\hat{\alpha_2} < \hat{\alpha_1}$	$\hat{\alpha_3} < 0$	$\hat{\alpha_4} < 0$ and $\hat{\alpha_4} > \hat{\alpha_3}$
Pattern 37	$\hat{\alpha_1} > 0$	$\hat{\alpha_2} > 0$ and $\hat{\alpha_2} \geq \hat{\alpha_1}$	$\hat{\alpha_3} < 0$	$\hat{\alpha_4} < 0$ and $\hat{\alpha_4} \leq \hat{\alpha_3}$
Pattern 38	$\hat{\alpha_1} > 0$	$\hat{\alpha_2} > 0$ and $\hat{\alpha_2} \geq \hat{\alpha_1}$	$\hat{\alpha_3} < 0$	$\hat{\alpha_4} < 0$ and $\hat{\alpha_4} > \hat{\alpha_3}$
Pattern 39	$\hat{\alpha_1} > 0$	$\hat{\alpha_2} > 0$ and $\hat{\alpha_2} \geq \hat{\alpha_1}$	$\hat{\alpha_3} < 0$	$\hat{\alpha_4} > 0$
Pattern 40	$\hat{\alpha_1} > 0$	$\hat{\alpha_2} > 0$ and $\hat{\alpha_2} < \hat{\alpha_1}$	$\hat{\alpha_3} < 0$	$\hat{\alpha_4} > 0$
Pattern 41	$\hat{\alpha_1} > 0$	$\hat{\alpha_2} > 0$ and $\hat{\alpha_2} < \hat{\alpha_1}$	$\hat{\alpha_3} > 0$ and $\hat{\alpha_3} > \hat{\alpha_2}$	$\hat{\alpha_4} < 0$
Pattern 42	$\hat{\alpha_1} > 0$	$\hat{\alpha_2} > 0$ and $\hat{\alpha_2} < \hat{\alpha_1}$	$\hat{\alpha_3} > 0$ and $\hat{\alpha_3} \leq \hat{\alpha_2}$	$\hat{\alpha_4} < 0$
Pattern 43	$\hat{\alpha_1} > 0$	$\hat{\alpha_2} > 0$ and $\hat{\alpha_2} \geq \hat{\alpha_1}$	$\hat{\alpha_3} > 0$ and $\hat{\alpha_3} < \hat{\alpha_2}$	$\hat{\alpha_4} < 0$
Pattern 44	$\hat{\alpha_1} > 0$	$\hat{\alpha_2} > 0$ and $\hat{\alpha_2} \geq \hat{\alpha_1}$	$\hat{\alpha_3} > 0$ and $\hat{\alpha_3} \geq \hat{\alpha_2}$	$\hat{\alpha_4} < 0$
Pattern 45	$\hat{\alpha_1} > 0$	$\hat{\alpha_2} < 0$	$\hat{\alpha_3} > 0$	$\hat{\alpha_4} > 0$ and $\hat{\alpha_4} \geq \hat{\alpha_3}$
Pattern 46	$\hat{\alpha_1} > 0$	$\hat{\alpha_2} < 0$	$\hat{\alpha_3} > 0$	$\hat{\alpha_4} > 0$ and $\hat{\alpha_4} < \hat{\alpha_3}$
Pattern 47	$\hat{\alpha_1} > 0$	$\hat{\alpha_2} > 0$ and $\hat{\alpha_2} \geq \hat{\alpha_1}$	$\hat{\alpha_3} > 0$ and $\hat{\alpha_3} \geq \hat{\alpha_2}$	$\hat{\alpha_4} > 0$ and $\hat{\alpha_4} < \hat{\alpha_3}$
Pattern 48	$\hat{\alpha_1} > 0$	$\hat{\alpha_2} > 0$ and $\hat{\alpha_2} \geq \hat{\alpha_1}$	$\hat{\alpha_3} > 0$ and $\hat{\alpha_3} < \hat{\alpha_2}$	$\hat{\alpha_4} > 0$ and $\hat{\alpha_4} \geq \hat{\alpha_3}$
Pattern 49	$\hat{\alpha_1} > 0$	$\hat{\alpha_2} > 0$ and $\hat{\alpha_2} \geq \hat{\alpha_1}$	$\hat{\alpha_3} > 0$ and $\hat{\alpha_3} \geq \hat{\alpha_2}$	$\hat{\alpha_4} > 0$ and $\hat{\alpha_4} \geq \hat{\alpha_3}$
Pattern 50	$\hat{\alpha_1} > 0$	$\hat{\alpha_2} > 0$ and $\hat{\alpha_2} \geq \hat{\alpha_1}$	$\hat{\alpha_3} > 0$ and $\hat{\alpha_3} < \hat{\alpha_2}$	$\hat{\alpha_4} > 0$ and $\hat{\alpha_4} < \hat{\alpha_3}$
Pattern 51	$\hat{\alpha_1} > 0$	$\hat{\alpha_2} > 0$ and $\hat{\alpha_2} < \hat{\alpha_1}$	$\hat{\alpha_3} > 0$ and $\hat{\alpha_3} > \hat{\alpha_2}$	$\hat{\alpha_4} > 0$ and $\hat{\alpha_4} \leq \hat{\alpha_3}$
Pattern 52	$\hat{\alpha_1} > 0$	$\hat{\alpha_2} > 0$ and $\hat{\alpha_2} < \hat{\alpha_1}$	$\hat{\alpha_3} > 0$ and $\hat{\alpha_3} \leq \hat{\alpha_2}$	$\hat{\alpha_4} > 0$ and $\hat{\alpha_4} \leq \hat{\alpha_3}$
Pattern 53	$\hat{\alpha_1} > 0$	$\hat{\alpha_2} > 0$ and $\hat{\alpha_2} < \hat{\alpha_1}$	$\hat{\alpha_3} > 0$ and $\hat{\alpha_3} > \hat{\alpha_2}$	$\hat{\alpha_4} > 0$ and $\hat{\alpha_4} > \hat{\alpha_3}$
Pattern 54	$\hat{\alpha_1} > 0$	$\hat{\alpha_2} > 0$ and $\hat{\alpha_2} < \hat{\alpha_1}$	$\hat{\alpha_3} > 0$ and $\hat{\alpha_3} \leq \hat{\alpha_2}$	$\hat{\alpha_4} > 0$ and $\hat{\alpha_4} > \hat{\alpha_3}$

Table .1: Table of the possible relative change slope patterns possible from our new definition of a Breakfast Canyon, with the first segment starting as a positive slope. These patterns correspond with Figure .1.

Bibliography

Funding for this work was partially supported by the National Science Foundation through NSF DMS-2015273. The findings and conclusions do not necessarily reflect the view of the funding agency.

[1] A., Akhila, B., Manjunatha, E., Keshamma (2022). Role of Honeybees in Biodiversity Conservation. *International Journal of Pharmacy and Biological Sciences*, 12(1).

[2] The Effect of Weather upon the Change in Weight of a Colony of Bees during the Honey Flow; United States Department of Agriculture: Washington, DC, USA, 1925; pp. 1–52.

[3] Holst, N., Meikle, W. (2018). Breakfast Canyon discovered in Honeybee hive weight curves. *Insects*, 9(4), 176.

[4] Insolia, L., Molinari, R., Rogers, S. R., Williams, G. R., Chiaromonte, F., Calovi, M. (2022). Honey bee colony loss linked to parasites, pesticides and extreme weather across the United States. *Scientific Reports*, 12(1).

[5] Karbassioon, A., Yearlsey, J., Dirilgen, T., Hodge, S., Stout, J. C., Stanley, D. A. (2023). Responses in Honeybee and bumblebee activity to changes in weather conditions. *Oecologia*, 201(3), 689–701.

[6] Khalifa, S. A., Elshafiey, E. H., Shetaia, A. A., El-Wahed, A. A., Algethami, A. F., Musharraf, S. G., AlAjmi, M. F., Zhao, C., Masry, S. H., Abdel-Daim, M. M., Halabi, M. F., Kai, G., Al Naggar, Y., Bishr, M., Diab, M. A., El-Seedi, H. R. (2021). Overview of bee pollination and its economic value for crop production. *Insects*, 12(8), 688.

[7] Kiran, M. S., Gündüz, M. (2014). The analysis of peculiar control parameters of artificial bee colony algorithm on the numerical optimization problems. *Journal of Computer and Communications*, 02(04), 127–136.

[8] Klein, A.-M., Vaissière, B. E., Cane, J. H., Steffan-Dewenter, I., Cunningham, S. A., Kremen, C., Tscharntke, T. (2006). Importance of pollinators in changing landscapes for world crops. *Proceedings of the Royal Society B: Biological Sciences*, 274(1608), 303–313.

[9] Klein, S., Pasquaretta, C., He, X. J., Perry, C., Søvik, E., Devaud, J.-M., Barron, A. B., Lihoreau, M. (2019). Honey bees increase their foraging performance and frequency of pollen trips through experience. *Scientific Reports*, 9(1).

[10] Meikle, W. G., Holst, N., Colin, T., Weiss, M., Carroll, M. J., McFrederick, Q. S., Barron, A. B. (2018). Using within-day hive weight changes to measure environmental effects on honey bee colonies. *PLOS ONE*, 13(5).

[11] Moritz, R.F., Kraus, F. B., Kryger, P., Crewe, R. M. (2007). The size of wild honeybee populations (apis mellifera) and its implications for the conservation of Honeybees. *Journal of Insect Conservation*, 11(4), 391–397.

[12] Nunes-Silva, P., Hrncir, M., Guimar aes, J. T., Arruda, H., Costa, L., Pessin, G., Siqueira, J. O., de Souza, P., Imperatriz-Fonseca, V. L. (2018). Applications of RFID technology on the study of bees. *Insectes Sociaux*, 66(1), 15–24.

[13] Quinlan, G. M., Isaacs, R., Otto, C. R., Smart, A. H., Milbrath, M. O. (2023). Association of excessive precipitation and agricultural land use with honey bee colony performance. *Landscape Ecology.*

[14] Sanchez-Bayo, F., Goka, K. (2014). Pesticide residues and bees – a risk assessment.*PLoS ONE*, 9(4).

[15] Schmickl, T., Crailsheim, K. (2004). Inner nest homeostasis in a changing environment with special emphasis on honey bee brood nursing and pollen supply. *Apidologie*, 35(3), 249–263.

[16] Seeley, T. D. (1982). Adaptive significance of the age polyethism schedule in Honeybee colonies. *Behavioral Ecology and Sociobiology*, 11(4)

[17] Seeley, T. D. (1995). *The wisdom of the hive: The Social Physiology of Honey Bee Colonies.* Acme Bookbinding.

[18] Schilling, R. L., Partzsch, L., Böttcher, B. (2021). *Brownian motion: An introduction to stochastic processes.* Walter de Gruyter.

[19] Steffan-Dewenter, I., Tscharntke, T. (2000). Resource overlap and possible competition between honey bees and wild bees in Central Europe. *Oecologia*, 122(2), 288–296.

[20] Streit, S., Bock, F., Pirk, C. W. W., Tautz, J. (2003). Automatic life-long monitoring of individual insect behaviour now possible.*Zoology*, 106(3), 169–171.

[21] Myerscough, M. R., Khoury, D. S., Ronzani, S., Barron, A. B. (2016). Why do hives die? using mathematics to solve the problem of honey bee colony collapse. *The Role and Importance of Mathematics in Innovation*, 35–50.

[22] Rosenkranz, P., Aumeier, P., Ziegelmann, B. (2010). Biology and control of Varroa destructor. *Journal of Invertebrate Pathology*, 103.

[23] vanEngelsdorp, D., Evans, J. D., Saegerman, C., Mullin, C., Haubruge, E., Nguyen, B. K., Frazier, M., Frazier, J., Cox-Foster, D., Chen, Y., Underwood, R., Tarpy, D. R., Pettis, J. S. (2009). Colony collapse disorder: A descriptive study.*PLoS ONE*, 4(8).

[24] Zaman, A., Dorin, A. (2023). A framework for better sensor-based Beehive Health Monitoring. *Computers and Electronics in Agriculture*, 210, 107906.