## Introduction Outline

*Preamble: Below is roughly the narrative direction I’m thinking for the introduction. Of course, a lot of details need to be filled in, but I try to outline the flow as best I can.*

The distribution and survival of invertebrate species are governed by a complex interplay of biotic and abiotic factors. For many insects, abiotic conditions such as temperature, humidity, and wind can be primary drivers of behavior, physiology, and habitat selection. Wind, in particular, can influence everything from dispersal and migration to foraging efficiency and predation risk.

The monarch butterfly (*Danaus plexippus*) presents an excellent system for studying these dynamics. Its diverse life history, which includes distinct breeding and overwintering phases, exposes it to a unique suite of environmental challenges. As a charismatic and threatened species, understanding the factors that limit monarch populations is not only of scientific interest but also critical for effective conservation.

The overwintering stage, when monarchs form dense aggregations in coastal California and central Mexico, is a period of immense physiological stress where abiotic factors are paramount. During this time, the "microclimate hypothesis" suggests that monarchs select groves that provide a specific suite of conditions necessary for survival: protection from freezing temperatures, access to moisture, dappled sunlight for thermal regulation, and shelter from wind.

Foundational work by Kingston Leong has been central to our understanding of monarch overwintering needs, specifically identifying wind as a key determinant of habitat suitability. Leong (2016) posits that winds exceeding 2 m/s are "disruptive," causing butterflies to be either physically dislodged from their roosts or induced to flee. Such disruption is considered detrimental as it can force the butterflies to expend critical energy reserves needed to survive the winter and may increase their exposure to predation. This assertion directly informs our choice of response variable: the change in monarch abundance at a roost.

Leong’s understanding of wind’s impact has been highly influential, forming the basis for management guidelines by organizations like the Xerces Society. However, his conclusions were largely inferred from comparing wind measurements at occupied roost trees versus unoccupied trees, rather than from direct observation of butterfly behavioral responses to wind. In contrast to Leong’s methods, this study uses direct, continuous time-series observation of behavioral responses within a single, occupied roost. To our knowledge, this relationship has never been directly and empirically tested.

Furthermore, to isolate the effect of wind, we must also account for other environmental drivers. Monarchs are ectotherms that rely on solar radiation to warm their bodies for flight. Exposure to sunlight is therefore a primary driver of activity and can trigger dispersal as individuals leave the roost to forage for nectar, seek water, or find mates. Therefore, to accurately test the disruptive wind hypothesis, we must statistically control for the confounding effects of sunlight exposure and ambient temperature. We include these as an interaction term because the effect of solar radiation on a butterfly’s body temperature is not independent of the ambient temperature. Direct sunlight provides the primary thermal energy for flight, but the amount of warming required to reach a flight threshold is lower on a warm day than on a cold one, making the effects of these two variables inherently interconnected.

While this 2 m/s threshold provides a critical starting point, the nature of wind is complex. Its disruptive force may be a function of not just a single speed but also its consistency (sustained wind) and turbulence (gustiness). Therefore, a comprehensive test of the disruptive wind hypothesis requires examining these different wind characteristics within the broader context of confounding effects.

## Hypothesis and Proposed Analysis

### Disruptive Wind Hypothesis

Based on the mechanistic framework established by Leong (2016), we hypothesize that disruptive wind events cause monarch butterflies to abandon their roosts. We predict that exposure to wind speeds exceeding the 2 m/s threshold will lead to a measurable decrease in local abundance such that winds exceeding the threshold will be negatively correlated with changes in monarch abundance at a roost.

### A Sequential Three-Step Analysis

To test this hypothesis, we propose a three-step analytical framework designed to first test the specific, proposed threshold (Step 1) and then, if necessary, explore the relationship more broadly (Steps 2 and 3). While I initially wanted to model the change in abundance directly, this approach creates a response variable with a complex and non-standard statistical distribution, making model interpretation unreliable.

#### Step 1: Direct Test of the 2 m/s Threshold

Our first step is a direct test of Leong’s hypothesis using Generalized Linear Mixed-Effects Models (GLMMs). To properly model our count-based abundance index and account for likely overdispersion, we will use a negative binomial distribution. The response variable will be the ‘abundance\_index‘ at time t, a metric representing the sum of conservative butterfly estimates within the sampled image. To account for temporal dependency, we will include the ‘abundance\_index‘ from the previous time step (t-1) as a covariate. We will create two predictor variables representing the number of minutes that sustained and gust wind speeds exceed the 2 m/s threshold and test them in separate models to avoid issues with collinearity.

% Load required library  
library(lme4)  
  
# Model 1: Test sustained wind threshold with Negative Binomial GLMM  
model\_sustained\_threshold <- glmer(  
 abundance\_index\_t ~ abundance\_index\_t\_minus\_1 +   
 sustained\_minutes\_above\_2ms +   
 sunlight\_exposure\_prop \* ambient\_temp +  
 (1 | view/labeler),  
 family = "nbinom2",  
 data = monarch\_data  
)  
  
# Model 2: Test gust wind threshold with Negative Binomial GLMM  
model\_gust\_threshold <- glmer(  
 abundance\_index\_t ~ abundance\_index\_t\_minus\_1 +   
 gust\_minutes\_above\_2ms +   
 sunlight\_exposure\_prop \* ambient\_temp +  
 (1 | view/labeler),  
 family = "nbinom2",  
 data = monarch\_data  
)

If either analysis reveals a negative relationship, it would provide strong empirical support for the 2 m/s threshold. If not, it would suggest the relationship is more complex, leading us to the next step.

#### Step 2: General Wind-Abundance Relationship Model

If the specific 2 m/s threshold is not a strong predictor, we will broaden our inquiry to test whether wind, in a more general sense, influences monarch abundance. While I initially wanted to model the change in abundance directly, this approach creates a response variable with a complex and non-standard statistical distribution, making model interpretation unreliable. To overcome this, I propose we use a more robust method by modeling the abundance\_index at each time step directly within a Generalized Linear Mixed-Effects Model (GLMM) framework, using a negative binomial distribution that is commonly used for count data.

A key challenge in analyzing time-series data is that the number of butterflies at any given moment is highly dependent on the number present just before. To properly account for this, our model will include the abundance from the previous time step (abundance\_index\_t\_minus\_1) as a predictor variable. By doing so, the model effectively asks: "Given the starting number of butterflies, what is the additional influence of our environmental predictors (e.g., wind, sun, temperature) on the abundance in the current time step?" This **method** allows us to statistically isolate the factors driving change between observations without creating a problematic response variable. It serves as more robust alternative for accounting autocorrelation than AR1.

To identify the most relevant aspect of wind, we will test a suite of continuous wind metrics, mean wind speed (sustained wind), maximum wind speed (peak gusts), and wind speed variance (gustiness), in separate models to avoid issues with collinearity. We will then use a model selection criterion, such as the Akaike Information Criterion (AICc), to compare these models. To ensure valid comparisons, the random effects structure will be held constant across all models. This entire analytical process will be repeated across multiple temporal scales (e.g., 30-minute, 4-hour, and daily) to investigate whether the disruptive effects of wind are due to immediate behavioral responses to acute events or cumulative responses to sustained conditions over longer durations. These scales were chosen to capture a range of potential biological responses, from short-term behavioral adjustments (30-minute) to responses to prolonged weather patterns (4-hour and daily), though the final scales may be adjusted based on data exploration.

# Example model for one wind metric (run separately for each)  
model\_max\_wind <- glmer(  
 abundance\_index\_t ~ abundance\_index\_t\_minus\_1 +   
 max\_wind\_speed +   
 sunlight\_exposure\_prop \* ambient\_temp +  
 (1 | view/labeler),  
 family = "nbinom2",  
 data = monarch\_data  
)

If this general model reveals that one or more wind metrics are significant predictors, we will proceed to the final step.

#### Step 3: Predictive Roost Abandonment Model

If we establish a significant relationship in Step 2, our final step will be to build a predictive roost abandonment model. Using a mixed-effects logistic regression model, we will predict the probability of a roost abandonment event (e.g., a drop in abundance exceeding a high threshold, such as >90%, to be finalized after data exploration), coded as a binary variable ‘roost\_abandoned‘. This model will use the most influential wind metric from Step 2 as the primary predictor.

% Load required library  
library(lme4)  
  
# Model: Predictive Hazard Analysis  
hazard\_model <- glmer(  
 roost\_abandoned ~ max\_wind\_speed +   
 sunlight\_exposure\_prop \* ambient\_temp,  
 family = binomial, # Specifies logistic regression  
 data = monarch\_data,  
 control = glmerControl(optimizer = "bobyqa")  
)

The goal is to produce a probabilistic curve that quantifies the risk of roost abandonment across a range of wind speeds, providing a nuanced tool for habitat managers.

## Dropping the Site Fidelity and Light Hypothesis

An initial version of this introduction included a hypothesis regarding the long-term loss of site fidelity. I propose to drop this hypothesis for a critical reason: our data cannot provide a convincing answer. Our methodology allows us to count butterflies, but not to track individuals. We cannot know if butterflies that appear after a dispersal event are the same individuals returning or new ones arriving. Answering the site fidelity question is probably better suited for a future project employing different methods. I prefer keeping this project clean and focused on answering the question of whether monarchs respond to strong winds.

Additionally, I also propose that we drop an explicit test of light mediated dispersal. The more I thought and read about potentially hypotheses, the more mired I became in thermal behavior, which this study is not well suited to address. I do think, however, that treating exposure to prolonged, direct light is a confounding variable that we need to address, and makes our analysis much stronger for incorporating it. If the data is compelling, perhaps we can include some graphs that describe the pattern of movement based on sunlight, but as of now, I can’t think of a defensible hypotheses we can answer with our data.

## Potential Ancillary Analysis: Rocket Launches

As a brief note, our dataset contains a unique, opportunistic record of monarch roost behavior during several rocket launches from nearby Vandenberg Space Force Base. If time permits, we could conduct a simple analysis to see if monarchs respond to the acoustic and vibrational disturbances. The method would be to compare the change in monarch abundance in the 30 minutes before and after a launch to changes during randomly selected 30-minute intervals from the study period. A null finding would itself be noteworthy. This could be a candidate for a short note publication.

Here is the summary table I presented for Jessica Griffiths regarding monarchs response to rocket launches:

| **Flight #** | **Date** | **Time** | **Boostback** | **Temp (F)** | **Monarch Change** | **Deployment(s)** | **Video(s)** |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 275 | 11/20/23 | 2:30:00 | No | 50 | No | SC3 | [SC3 video 1](https://youtu.be/EjNyWgDMnaA?si=tF78Lck0BBfF1kzb&t=21) |
| 278 | 12/1/23 | 10:19:00 | Yes | 66 | No | SC3 | [SC3 video 2](https://youtu.be/EjNyWgDMnaA?si=sDoxxwUsjJZVCXST&t=561) |
| 281 | 12/8/23 | 0:03:00 | No | 51 | No | SC4 | [SC4 video](https://youtu.be/RW91J3HMm_4?si=LM_TRjQAb_pA5lFN&t=104) |
| 284 | 12/24/23 | 5:11:00 | Yes | 46 | No | SC6 | [SC6 video 1](https://youtu.be/MqmH5L3lXbk?si=6g8voV1f43D4lI5Y&t=207) |
| 286 | 1/2/24 | 19:44:00 | No | 53 | No | SC6 | [SC6 video 2](https://youtu.be/MqmH5L3lXbk?si=aT3sviJN7YJ9i5bv&t=437) |
| 289 | 1/14/24 | 0:59:00 | No | 50 | No | SC8, SC10 | [SC8 video](https://youtu.be/AOWrySdh1z8?si=Pak3IK64mY7OlAMI&t=200) [SC10 video](https://youtu.be/0zLjm9PA6YM?si=Bokwr4KQj_jWIxWe&t=224) |
| 292 | 1/23/24 | 16:35:00 | No | 55 | No | SC10, SC8 (behind year) | [SC10 video](https://youtu.be/0zLjm9PA6YM?si=Ve9XAPoi9p32KaN0&t=431) [SC8 video (BY)](https://youtu.be/AOWrySdh1z8?si=744XzLu4L5U3Pjx7&t=431) |
| 294 | 1/28/24 | 21:57:00 | No | 59 | No | SC12, SC10 | [SC12 video](https://youtu.be/4uLZIZqlgPE?si=jwAqUIh4ifB_exz-&t=32) [SC10 video](https://youtu.be/0zLjm9PA6YM?si=JJND2CcZOMcbKwkF&t=552) |