6-21-2022 2:00 PM

Interdisciplinary Knowledge Exchange in Statistics with Applications in Fire Science and Statistical Education

Chelsea Uggenti, *The University of Western Ontario*

Supervisor: Woolford, Douglas G., *The University of Western Ontario*
Co-Supervisor: Dean, Charmaine B., *University of Waterloo*
A thesis submitted in partial fulfillment of the requirements for the Doctor of Philosophy degree in Statistics and Actuarial Sciences

Abstract

This thesis considers and articulates principles of interdisciplinary knowledge exchange in two contexts: the development of novel techniques for the study of wildland fire lifetimes; and, to understand improvements in the student training environment focusing on graduate teaching assistants, developing a training program on active learning for graduate teaching assistants in the mathematical and statistical sciences.

Wildland fire science is an area of research that requires interdisciplinary expertise to advance its body of knowledge. Wildland fires that are suppressed have a “lifetime” that consists of several sequential phases, including what are called *detection* and *action* phases. The interconnectedness of these phases is often overlooked when studying fire responses, and we develop methods to fill that gap in this thesis. In particular, we consider such a framework for the analysis of fire data from the Sioux Lookout District in northwestern Ontario. Multi-state modelling and joint frailty modelling techniques are employed. Comparisons of different frailty distributions and random effect forms are considered, and a simulation study is performed to highlight the advantages of a flexible model form for the joint frailty models. Using the joint frailty models, we find that fires with longer detection phases are associated with longer action phases, and that the action phase lengths may be increasing over time. Collaboration with fire scientists throughout the development of this work was critical and is especially important for ensuring the impact of it at fire management agencies.

The importance of collaboration in statistics is emphasized in how education in this field is conducted. A workshop on active learning techniques, which aid in the exchange of knowledge between students and instructors, was developed for graduate teaching assistants at the University of Western Ontario. A survey study of graduate teaching assistant perceptions about active learning before and after participating in a workshop on active learning in mathematical and statistical sciences was performed. Learnings from this study are discussed.

Keywords: Fire Lifetime, Graduate Teaching Assistant Training, Joint Frailty Modelling, Multi-State Modelling, Statistical Education, Survival Analysis, Wildland Fire

Summary for Lay Audience

This thesis applies the principles of knowledge exchange — a "push and pull" of information — in the development of novel techniques for the study of wildland fire lifetimes and the development of a training program to enhance the statistical education of graduate teaching assistants.

Our first study focuses on the lifetimes of suppressed wildland fires that have several sequential phases. We are interested in characterizing what drives these phases as well as understanding how the time in earlier phases may impact the latter portion of a fire's lifetime. We consider the detection phase, consisting of the time from the ignition to the report of a fire, and the action phase, consisting of the time from report to being declared "under control". Fires with a longer detection phase can have a longer than anticipated action phase, when compared to fires that were reported quicker. This makes sense because if it takes longer to find and report a fire after it is ignited then it may grow larger and could possibly take longer to bring under control. By explicitly linking the two phases, we can identify how they are connected or interact with one another. Knowledge exchange was used throughout the entire process of studying wildland fire lifetimes by attending interdisciplinary conferences, engaging with interdisciplinary researchers, and collaborating with fire scientists, to name a few approaches.

Our second study moves away from wildland fire and into statistical education since the importance of collaboration in statistics is emphasized in how education in this field is conducted. The training program developed for graduate teaching assistants consists of a workshop on active learning, which are techniques that aid in the exchange of knowledge between students and instructors (or graduate teaching assistants). We performed a survey study at the University of Western Ontario to examine graduate teaching assistant perceptions about active learning before and after participating in a workshop on active learning in mathematical and statistical sciences. Learnings from this study are discussed.

Co-Authorship Statement

Paper Title: Characterizing and Linking Two Phases of Wildland Fire Lifetimes in Ontario

Publication: In preparation.

List of Authors: Chelsea Uggenti, C.B. Dean, Douglas Woolford, and Colin McFayden

Description: Dr. Dean and Dr. Woolford suggested the use of multi-state models and joint frailty models for the two lifetime phases. The structure of the model was inspired by conversations with Dr. Dean and Dr. Woolford, along with the works of Dr. Nathoo and Dr. Xi (former students of Dr. Dean). Mr. McFayden provided feedback and suggestions on several items, including the understanding of key fire events, important variables to consider when modelling, and insights gained from the results. Dr. Dean suggested a simulation study to analyze the performance of the model form. All data manipulations, analysis, diagnostics, simulations, and visualizations were performed by me. Chapters 3 and 4 will be turned into a paper to be submitted to the peer-reviewed literature, with me acting as the lead author and the others as co-authors.

Paper Title: Characterizing the Detection and Action Phases for Suppressed Wildland Fires in the Sioux Lookout District

Publication: In preparation. Internal document for the Ontario Ministry of Northern Development, Mines, Natural Resources, and Forestry.

List of Authors: Chelsea Uggenti, Colin McFayden, Douglas Woolford, and C.B. Dean

Description: This wildland fire science communiqué serves as an internal document for Ministry personnel on some of the findings related to this thesis. The structure of the communiqué was inspired by conversations with Mr. McFayden and Dr. Woolford along with a previous communiqué by Trish Ogen (a former employee of the Ministry and Research Assistant of Dr. Woolford). Mr. McFayden was instrumental in assisting with the clarity of the communiqué to ensure effective knowledge exchange with non-statisticians within the target audience. All analyses and visualizations were performed by me. I was the lead author of this communiqué, with input and suggested revisions from the others as co-authors.

Paper Title: Benefits, Limitations, and Practical Strategies for Implementing Active Learning Activities in Undergraduate University Mathematics and Statistics Courses: A Graduate Teaching Assistant Training and Development Program
Publication: In preparation.
List of Authors: Chelsea Uggenti, C.B. Dean, and Douglas Woolford
Description: This publication discusses a workshop on active learning that was developed as part of this thesis. The initial version of this workshop was created in 2018 and a final version was created in March 2021. All ideas and methods used were developed by me. The paper will be written by me as lead author, with input and suggested revisions from the others as co-authors.

Paper Title: Investigating Graduate Teaching Assistant Training and Development in Western University's School of Mathematical and Statistical Sciences
Publication: In preparation.
List of Authors: Chelsea Uggenti, C.B. Dean, and Douglas Woolford
Description: The idea for this study was suggested by me. Dr. Woolford provided valuable comments on the design of the study and support for its implementation. The Non-Medical Research Ethics Board application was written by me with input and assistance from Dr. Woolford. All data manipulation, analysis, and visualizations were performed by me. The paper will be written by me as lead author, with input and suggested revisions from the others as co-authors.

Acknowlegements

Thank you to each person who has helped me on this journey.

Most notably I wish to acknowledge the support, encouragement, contributions, and training from my co-supervisors. Dr. Woolford opened my eyes to the world of academia and research as a second-year undergraduate student. He has offered endless academic, professional, and personal guidance ever since. Dr. Dean asked thought-provoking questions and shared deep insights that have helped me to produce better quality work by seeing the "big picture". Thank you both for your one-of-a-kind mentorship.

Thank you to the examining committee for providing insightful remarks and questions that have improved this thesis. These reviewers are Dr. Bethany White, Dr. Pauline Barmby, Dr. Camila de Souza, and Dr. Simon Bonner.

I would like to thank Colin McFayden, Dr. Elizabeth Renouf, and Dr. Den Boychuk for their many helpful conversations, feedback, and suggestions. I would also like to thank Dr. Meg Krawchuk and her Landscape Fire and Conservation Science Research Group members for offering me the opportunity to gain fieldwork experience and for the eye-opening discussions about fire and fire ecology. As well, I would like to thank Dr. Lisa Aikman and Dr. Ken Meadows from Western's Centre for Teaching and Learning for their assistance and guidance, along with the School of Mathematical and Statistical Sciences for their support of our study. Thank you to the Ontario Ministry of Northern Development, Mines, Natural Resources and Forestry for providing the wildland fire datasets and for their active engagement in this research.

Thank you to past and present lab members who have played a significant role in my research and development, including Devan Becker, John Thompson, Dexen Li, Kevin Granville, and Johanna de Haan Ward.

Support provided by the following sources is gratefully acknowledged: the Natural Sciences and Engineering Research Council of Canada (NSERC); the Canadian Statistical

Sciences Institute (CANSSI); the Institute for Catastrophic Loss Reduction (ICLR); the TD Bank group for their generous donation to support data analytics in Western's Faculty of Science, including the funds used to create the TD Professor in Data Analytics that helped partially support my Ph.D. studies; the University of Western Ontario and its Science International Engagment Fund, Centre for Environment and Sustainability, and Department of Statistical and Actuarial Sciences.

And, most importantly, my family and friends who are the foundation of my life — all of you have supported, encouraged, and celebrated my passions. Thank you from the bottom of my heart. Nicholas Sakellis joined me on this adventure and I am ecstatic to face our future adventures together. My mother has always been and will always be my loudest cheerleader. Thank you Mom for modelling a strong work ethic, imparting words of wisdom, fostering my dreams, and believing in me when I forget to believe in myself.

Dedicated to my mother. The strongest woman I know.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The field of statistics is both intriguing and fundamental to scientific inquiry, reaching across many other disciplines. A critical aspect of the interdisciplinary nature of statistics is how information is exchanged between those within the field and those outside of it. The exchanging of knowledge is often complex, dependent on the people involved and the environment created. A simple and naive question we, as statisticians, may ask is whether we should push statistical information onto others or if the exchange of information should follow more of a push-and-pull relationship? For example, a statistician may be viewed solely as technical support for a project with a lower expectation of knowledge exchange or may be asked to be an active collaborator on a project with a higher expectation of knowledge exchange. A supportive working environment requires members to actively listen, constructively communicate, openly share information, and collaboratively solve problems. Members that are flexible, reliable and dedicated to the team are also essential to fostering genuine knowledge exchange. This thesis considers the knowledge exchange of two main topics developed within an interdisciplinary learning environment (Wachowicz and Chrisman, 2012): statistics within the fields of wildland fire science and statistical education.

1.1 Fire Science

In a world of changing climate, we see its effects on wildland fire activity first hand. Media and news reporters cover stories of wildland fires burning for weeks in Canada, USA, Australia, and other places around the globe that destroy residences and high-value infrastructure, and threaten lives.

Fire management agencies spend close to one billion dollars a year on average to fight and suppress wildland fires in Canada (Natural Resources Canada, 2021a; Stocks and Martell, 2016). Taylor et al. (2013) argued that the field of statistical science can help by improving wildland fire predictions that could be used in decision-support tools by fire management agencies. Research on wildland fire behaviour, cost, and resource allocation is a necessity because it can help reinforce fire management best practices or call attention to any areas in need of improvement.

Our research in the wildland fire science domain focuses on several key phases of a fire that make up the total lifetime of a wildland fire. For instance, the detection phase encompasses the start of the fire to the time when it's reported, whereas the action phase involves the time from when it's reported to when it is brought under control. Past research by Morin et al. (2015) modelled a single phase of a fire's lifetime, namely the time until a fire is classified by a fire management agency as being under control from ignition, called the control time. Sun (2013), Morin et al. (2019), and Xi et al. (2020) also performed research on fire lifetimes where only one phase of wildland fire lifetimes was considered. By only considering a single phase, we have a crucial gap in understanding such lifetimes, namely understanding the fire's evolution over several phases and the interconnectedness of the phase lengths, if any.

In this dissertation, we consider how the distributions of time in the various phases of a fire, as it progresses from ignition to being under control, are connected and how different factors affect the time spent in these phases differently. We investigate the lifetime distributions of the various phases of wildland fires from a study area consisting of a fire management district located in Ontario's Northwestern Fire Region using multi-state modelling and joint frailty modelling techniques. These techniques provide a mechanism for modelling how the preceding phase(s) may affect the subsequent phase(s).

Multi-state models are used to characterize the evolution of individuals through a series of states (or phases) until they reach an endpoint (Cook and Lawless, 2018). In the fire context, multi-state models are used to represent the progression of a fire through its various phases. The phases can also be jointly modelled, where each phase represents a different outcome in the model. Joint frailty models are employed to explicitly link the fire lifetime phases to see how they interact with one another, rather than modelling each phase separately. Joint models have been applied to other topics in the fire context. For example, they have been used to model the duration and size of fires (Xi et al., 2019, 2021) since it is often the case that these responses are correlated.

We utilize the principles of knowledge exchange throughout the entire process of studying the phases of wildland fire lifetimes. Considerations for stakeholders and end-users of this research were fostered by several experiential learning opportunities, including attending interdisciplinary conferences, communicating with researchers in fire science and ecology and with fire management practitioners, and collaborating with a Fire Science Specialist from Ontario's Aviation, Forest Fire and Emergency Services branch of the Ontario Ministry of Nothern Development, Mines, Natural Resources and Forestry (MNDMNRF). The culmination of the latter resulted in a communiqué (i.e., a brief, concise, results-focused summary) of our research on wildland fire lifetime phases written for MNDMNRF members in an accessible and non-statistical way.

1.2 Statistical Education

A second key element of study in this thesis relates to statistical education. Many graduate teaching assistants (GTAs) begin their roles with little or no prior teaching experience.

Yet they play an essential role in undergraduate student learning and assessment, and are often a first point of contact for undergraduate students. As students often see GTAs as less intimidating figures than their professors, GTAs have great potential to engage and inspire the future scholars from their disciplines (Dimitrov et al., 2013).

Both the responsibilities that science graduate teaching assistants undertake and the volume of science undergraduate courses being taught at research universities by contract lecturers who primarily only have GTA experience are increasing (see Gardner and Jones (2011) and references therein). Teaching-related training of GTAs in the statistical and mathematical sciences is often limited, informal, or under-developed, and typically arises from the reflection of the experience of being students themselves or "on the job" trial-and-error experiences (Gelman, 2005; Gardner and Jones, 2011). This issue amplifies when contract lecturers or professors who only have GTA teaching experience instruct introductory science, technology, engineering and mathematics (STEM) courses since they may lack the pedagogical skills to teach these courses effectively (Crowe, 2019).

There are various reasons why STEM GTAs need instructor training. Gelman (2005) notes that it may be hard for them to relate to the various types of learners in a course since graduate students are often top performers in similar environments. They are also inclined to use traditional lecture-style techniques because such approaches are familiar to them, and they may have developed rigid, deeply-held beliefs about teaching (Justice et al., 2017). Statistics GTAs often resist employing active learning techniques or participatory activities in their tutorials or lectures due to anxieties that they will not have time to cover what has been identified as *important* material (Gelman, 2005). Nevertheless, such techniques as active learning are vital to learning in statistics as it embodies principles of collaboration in teaching that are fundamental to statistics as an interdisciplinary science.

Gardner and Jones (2011) highlight the many challenges that GTAs may face regarding their "pedagogical preparedness". For instance, GTAs often feel overwhelmed with all the demands placed on them and their time, leading to a feeling of self-preservation reflected in their often narrow and restricted list of priorities. These pressures, combined with teaching development viewed as a low priority by most graduate students, cause them to develop their researcher identity at the expense of their instructor identity. Gardner and Jones (2011) also note that although knowing the content knowledge is important for effective teaching, it is problematic that neither the GTAs, nor many institutions and disciplines, prioritize pedagogical training as a requirement. Instead, the most consistent support that GTAs typically receive and utilize regarding their teaching comes from their peers and fellow GTAs.

The Scholarship of Teaching and Learning (SoTL) is a movement of scholarly thought and action that draws on the connected relationship between teaching and learning at the post-secondary level. It differs from traditional research since it is defined as "the systematic study of teaching and learning, using established or validated criteria of scholarship, to understand how teaching (beliefs, behaviours, attitudes, and values) can maximize learning, and/or develop a more accurate understanding of learning, resulting in products that are publicly shared for critique and use by an appropriate community" (Potter and Kustra, 2011, p. 2). Essentially, it is research dedicated to teaching and learning that can span across a multitude of disciplines.

Our SoTL research seeks to answer the following research question: **How does participation in a discipline-specific teaching development program on active learning for Graduate Teaching Assistants (GTA) in mathematics and statistics, offered by their School of Mathematical and Statistical Sciences, impact their perceptions of teaching?** We developed a 1.75-hour long workshop on active learning techniques for GTAs in the School of Mathematical and Statistical Sciences at the University of Western Ontario. Active learning is the embodiment of knowledge exchange since it requires that both students and instructors (or GTAs) actively engage with the material, more so on the students who are undertaking the active learning activity than the instructor who plans and facilitates it. These activities must take place within a supportive learning environment that helps to foster this exchange of knowledge. We employ a survey study design where participants were asked to attend the workshop and respond to pre-post survey questionnaires related to active learning and the workshop.

1.3 Intended Audience and Dissertation Outline

This dissertation is intended for multiple audiences. The fire lifetime modelling work may be of interest to statisticians who study time-to-event or lifetime data, and fire scientists or fire management practitioners interested in understanding wildland fire lifetimes. Whereas the SoTL work on GTA training and development is intended for educators (i.e., lecturers, supervisors, departments, etc.) within the fields of statistics and mathematics who are concerned with and care about the training and development that GTAs receive during their academic careers.

We utilize the statistical software R (R Core Team, 2021) to clean and wrangle data, create data visualizations, and perform analyses. Throughout the thesis we provide information on the R packages employed in our research.

The chapter structure of the dissertation is:

Chapter 2: A literature review of relevant theory.

Chapter 3: The characterization of two phases of wildland fire lifetimes using multi-state modelling techniques.

Chapter 4: The linking of those two phases with joint frailty models and evaluating the efficacy of the preferred models with both diagnostics and a simulation study.

Chapter 5: Personal reflections on the process of exchanging the fire science knowledge gained from the previous chapters.

Chapter 6: The SoTL study investigating the training and development of GTAs.

Chapter 7: A discussion on future work to close the dissertation.

Chapter 2

Background Theory

This chapter presents an overview of the key topics used in our work, including survival analysis, Cox proportional hazards models, multi-state models, mixed effects models, frailty models, and Bayesian methods and evaluation, along with the theories of knowledge exchange and active learning.

2.1 Survival Analysis

A *lifetime* is defined to be the time from a specific starting point until some well-defined event occurs, not necessarily the end of a life. Lifetime data is typically incomplete since we cannot always wait for the event to occur during a study, or we may start observing after the starting point that measures the lifetime, or observations may not be made continuously. A lifetime is *censored* if it is only known that it lies within some interval. A lifetime is *truncated* if its value is beyond the observation boundary. Censoring and truncation therefore commonly occur when dealing with lifetime (or time-to-event) data.

For a strictly positive continuous lifetime random variable T, the following functions are used to characterize its distribution. The *survival function*, $S(t) = P(T > t)$, is a nonincreasing function bounded in $[0, 1]$ that gives the probability that the lifetime exceeds t, where $t > 0$. The *hazard function*

$$h(t) = \frac{-\frac{d}{dt}S(t)}{S(t)},$$

represents the instantaneous rate of occurrence of the event for which the lifetime is measured at each point in time given that the event has not yet occurred. It is also known as the force of mortality or the failure rate.

The Kaplan-Meier estimator is a non-parametric statistic used to estimate the survival function from lifetime data. It is defined as

$$\widehat{S}(t) = \prod_{i:t_i \leq t} \left(1 - \frac{d_i}{n_i}\right),$$

where i indexes the unique times, t_i, when at least one event occurred before time t, d_i is the number of events (e.g., deaths) that happened at time t_i, and n_i is the number of *individuals known to have survived* (i.e., have not yet had an event or been censored) up to time t_i.

2.2 Cox Proportional Hazards Models

Cox proportional hazards (PH) models are commonly used for modelling the relationship between the predictors and the survival outcome (Cox, 1972). Let $x_{ij}(t)$ be the jth predictor of the ith person or observation at time t, where $i = 1, \ldots, n$ and $j = 1, \ldots, p$. Then we use the $p \times 1$ column vector $\boldsymbol{x}_i = (x_{i1}, \ldots, x_{ip})'$ for time-fixed predictors and $\boldsymbol{x}_i(t) = (x_{i1}(t), \ldots, x_{ip}(t))'$ for time-varying predictors, where some or all may be time-varying. If none are time-varying then $\boldsymbol{x}_i(t) = \boldsymbol{x}_i$. For observation i, the hazard function for the Cox PH model is defined as

$$h_i(t) = h_0(t) \exp(\boldsymbol{x}_i'(t)\boldsymbol{\beta}), \tag{2.1}$$

where h_0 is the unspecified baseline hazard function (which is the hazard corresponding to $\boldsymbol{x}_i(t)$ being the zero vector) and $\boldsymbol{\beta}$ is a $p \times 1$ column vector of regression coefficients. For two observations i and k, the ratio of their hazard functions with time-fixed predictors has the form

$$\frac{h_i(t)}{h_k(t)} = \frac{h_0(t)\exp(\boldsymbol{x}_i'\boldsymbol{\beta})}{h_0(t)\exp(\boldsymbol{x}_k'\boldsymbol{\beta})} = \frac{\exp(\boldsymbol{x}_i'\boldsymbol{\beta})}{\exp(\boldsymbol{x}_k'\boldsymbol{\beta})},$$

which is constant over time. Thus, the model is known as the *proportional hazards* model.

In general, the hazard ratio (HR), $\exp(\boldsymbol{\beta}) = (\exp(\beta_1), \exp(\beta_2), \ldots, \exp(\beta_p))'$, measures the effect of the predictor(s) on the hazard compared to the baseline hazard. If the jth element of the HR is greater than one then as the value of x_{ij} increases there is an increase in the hazard, resulting in a decreased chance of survival. Conversely, if the jth element of the HR is less than one then as the value of x_{ij} increases there is a decrease in the hazard, resulting in an increased chance of survival. If the jth element of the HR is equal to one then there is no effect. The `survival` package (Therneau, 2020) has several functions that fit these models (e.g., the `coxph` function).

Estimation of $\boldsymbol{\beta}$ is based on the partial likelihood function introduced by Cox (1972). Let t_i be the survival time associated with the ith individual, c_i be the fixed censoring time, $y_i = \min(t_i, c_i)$, and let the event indicator be

$$\nu_i = \begin{cases} 1, \text{ if } t_i \leq c_i \\ 0, \text{ if } t_i > c_i \end{cases} .$$

For continuous lifetime data with no ties the partial likelihood function has the form

$$PL(\boldsymbol{\beta}) = \prod_{i=1}^{n} \left(\frac{\exp(\boldsymbol{x}_i'\boldsymbol{\beta})}{\sum_{r \in \mathcal{R}_i} \exp(\boldsymbol{x}_r'\boldsymbol{\beta})} \right)^{\nu_i},$$

where $\mathcal{R}_i$ is the set of individuals who are “at risk” for the event at a time just prior to $y_{(i)}$, called the *risk set*, and $y_{(i)}$ is the ith ordered survival time. However, real data sets often contain tied events as lifetime data. Therneau and Grambsch (2000) discuss four common methods for handing ties: Breslow approximation, Efron approximation, exact partial likelihood, and averaged likelihood. If the data contains no ties then all these methods are equivalent.

2.2.1 Baseline Hazard Functions

For a *parametric* Cox PH model we assume a parametric function for the baseline hazard, $h_0(t)$. Two popular choices are:

1. $h_0(t) = \lambda \rho t^{\rho-1}$, with $\lambda > 0$, $\rho > 0$, or
2. $h_0(t) = \lambda \exp(\gamma t)$, with $\lambda > 0$, $\gamma \in \mathbb{R}$.

The first choice results in event times or lifetimes that have a Weibull distribution whereas the second choice results in a Gompertz distribution. Table 1.13 in Duchateau and Janssen (2007) outlines the hazard function, density function, and survival function for several distributions, including these two choices.

One popular *semiparametric* Cox PH model assumes a piecewise exponential baseline hazard function. This specification provides a very flexible framework for modelling the baseline hazard rate whose shape might not be as simple as those represented by a Weibull or Gompertz distribution. The time axis is partitioned into K prespecified intervals where

$$I_{(a_{k-1},a_k]}(t) = \begin{cases} 1, & t \in (a_{k-1}, a_k] \\ 0, & \text{otherwise} \end{cases}, \tag{2.2}$$

for $k = 1, \ldots, K$, where $0 = a_0 < a_1 < \cdots < a_K < \infty$, a_K is the last survival/censored time, and we assume that the baseline hazard is constant within the intervals (Ibrahim et al., 2001).

For more details on lifetime data, survival analysis and Cox PH models see Lawless (2011), Kleinbaum and Klein (2010), and Therneau and Grambsch (2000).

2.3 Multi-State Models

The following section on multi-state models, along with Chapter 2.4, use notation and writing adopted from Cook and Lawless (2018).

Life history data (or lifetime/time-to-event data) is implicitly longitudinal because it contains information about events or other outcomes observed over time. Let $N(t)$ be the number of events occurring up to time t. Then $\{N(t), t \geq 0\}$ is called a counting process. When there are $R \geq 2$ types of events, we let $N_r(t)$ denote the number of events of type r over the time interval $(0, t]$, where $r = 1, 2, ..., R$. Note that processes $\{N_r(t), t \geq 0\}$ are called counting processes of type r events. There are two features of interest for counting processes that we will focus on: the number of events that occurred over specific time periods and the lengths of time between specific events.

Now let $Z(t)$ denote the state of the process at time t, where there are $1, 2, ..., K$ mutually exclusive states. The researcher is free to define states in the way that best addresses their research question. In this case, the two features of interest are the probability of moving from one state to another and the duration of time spent in specific states.

Models for events and their counting processes in *continuous time* are specified through intensity functions. Let the σ-algebra $\mathcal{H}(t) = \{N(s), 0 \leq s \leq t\}$ denote the history for all events of $[0, t]$[1]. Then the intensity function for events of type r is

$$\lambda_r(t|\mathcal{H}(t^-)) = \lim_{\Delta t \searrow 0} \frac{P(\Delta N_r(t) = 1 \mid \mathcal{H}(t^-))}{\Delta t}, \ t \geq 0, \tag{2.3}$$

where $\Delta N_r(t) = N_r(t + \Delta t) - N_r(t)$ is the difference between the counting processes of type r events over a short time period Δt, $\mathcal{H}(t^-)$ represents the history of states that are occupied over $[0, t)$, and $\mathcal{H}(0^-) = \emptyset$. For a continuous-time process we assume two or more events cannot occur at the same time resulting in the intensity functions for

[1] A σ-algebra on a set $\mathcal{X}$ is a collection, $\Sigma^{\mathcal{H}}$, of subsets of $\mathcal{X}$ that includes $\mathcal{X}$, is closed under complement, and is closed under countable unions.

$r = 1, ..., R$ fully specifying the multivariate event process.

Multi-state models in *continuous time* with state space $\{1, 2, ..., K\}$ are formulated by specifying intensity functions, similar to (2.3), for allowable transitions between states. The **transition intensity functions** are

$$\lambda_{kl}(t|\mathcal{H}(t^-)) = \lim_{\Delta t \searrow 0} \frac{P[Z(t + \Delta t) = l \mid Z(t) = k, \mathcal{H}(t^-))]}{\Delta t}, \ k \neq l \tag{2.4}$$

representing the instantaneous rate of progression to state l conditionally on occupying state k. Here, we let $\mathcal{H}(t) = \{Z(s), 0 \leq s \leq t\}$. A multi-state model is a Markov process if $\lambda_{kl}(t|\mathcal{H}(t^-)) = \lambda_{kl}(t)$ since the transition intensity function depends on the history of the process only through the current state. Multi-state models can be represented as counting processes by expressing the types of transitions as different types of events, assuming only one event or 'transition' can occur at a given instant.

Intensities can also be specified as functions of predictors. For fixed predictors, $\mathcal{H}(t) = \{Z(s), 0 \leq s \leq t; \mathbf{X}\}$ where $\mathbf{X}$ is a fixed predictor matrix with rows $\boldsymbol{x}_i^{'}$, $i = 1, \ldots, n$. For individual i, the Markov model where predictors act multiplicatively on the intensity has the general form

$$\lambda_{i,kl}(t|\mathcal{H}(t^-)) = \lambda_{i,kl}(t|\boldsymbol{x}_i) = \lambda_{kl0}(t) g(\boldsymbol{x}_i; \boldsymbol{\beta}_{kl}),$$

where $\boldsymbol{x}_i$ are the predictors, $\boldsymbol{\beta}_{kl}$ the regression coefficients and $g(\boldsymbol{x}_i; \boldsymbol{\beta}_{kl}) \geq 0$.

A common choice for the function $g(\cdot)$ is $g(\boldsymbol{x}_i; \boldsymbol{\beta}_{kl}) = \exp(\boldsymbol{x}_i^{'} \boldsymbol{\beta}_{kl})$ which follows a similar framework to the Cox PH model (2.1). In this case $\lambda_{kl0}(t)$ are called **baseline intensities** which apply when $\boldsymbol{x}_i = \mathbf{0}$.

For individual i, the Markov model where predictors act additively on the intensity has the form

$$\lambda_{i,kl}(t|\mathcal{H}(t^-)) = \lambda_{kl0}(t) + g(\boldsymbol{x}_i; \boldsymbol{\beta}_{kl}),$$

where one must constrain the model components so the intensity is non-negative. Choosing $g(\boldsymbol{x}_i, t; \boldsymbol{\beta}_{kl}) = \boldsymbol{x}_i^{'} \boldsymbol{\beta}_{kl}(t)$ allows for time-dependent regression coefficients (Meira-Machado et al., 2009) which are discussed further in Chapter 3.

2.3.1 Features of Interest

The two main features of interest are **transition intensities** and **transition probabilities**. Transition intensities (denoted by λ_{kl}) "describe the instantaneous risk of a change in the process by specifying how the probability of a transition occurring over a short time interval depends on the process history up to that time" (Cook and Lawless, 2018, p. 9).

Transition probabilities are denoted by

$$P_{kl}(s,t|\mathcal{H}(s^-)) = P[Z(t) = l \mid Z(s) = k, \mathcal{H}(s^-)],$$

for $k, l \in \{1, ..., K\}$ and $s \leq t$. When individuals must be in state 1 at $s = 0$, $P_{1l}(0,t|\mathcal{H}(0))$ for $t > 0$ and $l \in \{1, ..., K\}$ are called *prevalence* or *occupancy* probability functions. These functions give the probability of moving from the initial state (in this case, state 1) to any other state, including state 1. Durations of sojourns in certain states or the time until a specific state is first entered can also be of interest.

2.3.2 Counting Process Definitions

It's often helpful to express data and models in terms of counting processes. Note the following definitions:

- $N_{kl}(t)$ = right-continuous function that counts the number of instantaneous transitions from k to l over $[0, t]$

- $\Delta N_{kl}(t) = N_{kl}(t + \Delta t) - N_{kl}(t)$ = number of k to l transitions over $[t, t + \Delta t]$

- $dN_{kl}(t) = \lim_{\Delta t \to 0} \Delta N_{kl}(t)$ indicates if k to l transitions occurred at t

$$\therefore dN_{kl}(t) = \begin{cases} 1 & \text{, if a } k \text{ to } l \text{ transition occurred at time } t \\ 0 & \text{, otherwise} \end{cases}$$

- Vector $dN_k(t) = (dN_{kl}(t), l \neq k, l = 1, \ldots, K)$ contains all the elements $dN_{kl}(t)$ for $l \neq k$ and hence all the information on whether a transition out of state k at time t occurred and the nature of the transition. Note that if $\sum dN_{kl}(t) = 1$ then a transition occurred and the non-zero element in $dN_k(t)$ tells the entered state.

- $N_k(t) = (N_{kl}(t), l \neq k, l = 1, \ldots, K)'$ where $N_{kl}(t) = \int_0^t dN_{kl}(s)$, gives the cumulative number and types of transitions out of state k of each type over $[0, t]$

- Full vector $N(t) = (N_1'(t), \ldots, N_K'(t))'$ records the nature and number of all transitions over [0,t]

Thus, $\{N(t), t \geq 0\}$ is another way of representing $\{Z(t), t \geq 0\}$, i.e. the counting process represents the multi-state framework.

The following conventions are utilized in the next section and are important to remember. $Y(t) = I(t \leq C)$ is a process that is under *observation* at time t where C is the right censoring time and thus $Y_k(t) = I(Z(t) = k)$ is an indicator function denoting that state k is occupied at time t. We use $\bar{Y}_k(t) = Y(t)Y_k(t^-)$[2] to indicate that a transition out of state k may be observed at time t.

[2] Note that this $\bar{Y}$ notation represents an *observation* rather than an average.

2.4 Analysis of Continuous Multi-State Models

2.4.1 Parametric Maximum Likelihood Estimation

Let θ be a vector of parameters whose elements correspond to the transition intensity functions $\lambda_{kl}(t|\mathcal{H}(t^-))$ for a multi-state process with K states. The vectors θ_{kl} $(l \neq k)$ parameterize λ_{kl} and, generally, $\theta_{kl} \neq \theta_{k'l'},\ k, l, k', l' = 1, \ldots, K,\ (k, l) \neq (k', l')$. Also assume that the observation of individual i begins at A_{i0} $(A_{i0} \geq 0)$ and stops at C_i $(C_i > A_{i0}),\ i = 1, \ldots, n$.

The **full likelihood** for $\theta = (\theta_{kl},\ k = 1, \ldots, K,\ l = 1, \ldots, K)$ is

$$L(\theta) = \prod_{k \neq l} L_{kl}(\theta_{kl}), \tag{2.5}$$

where

$$L_{kl}(\theta_{kl}) = \prod_{i=1}^{n} \left\{ \prod_{t_{ir} \in \mathcal{D}_{ikl}} \lambda_{kl}(t_{ir}|\mathcal{H}(t_{ir}^-); \theta_{kl}) \times \exp\left(- \int_0^\infty \bar{Y}_{ik}(u) \lambda_{kl}(u|\mathcal{H}_i(u^-); \theta_{kl}) du \right) \right\}, \tag{2.6}$$

with $i = 1, \ldots, n$ sampled independent individuals, $\bar{Y}_{ik}(u) = I(A_{i0} \leq u \leq C_i) I(Z_i(u^-) = k)$, and $\mathcal{D}_{ikl}$ is the set of distinct times t_{ir} at which individual i makes an observed $k \to l$ transition.

The full likelihood is an extension of the likelihood contribution for a single individual. Chapter 2.2 in Cook and Lawless (2018) provides details of how the single likelihood is developed using product integration and sample path probabilities. The time interval of an observed individual is partitioned so the likelihood contributions within the sub-intervals are considered, the products of these likelihood contributions are taken over the partition, and then the limit as the number of sub-intervals goes to infinity is taken.

The partial log-likelihood $l_{kl}(\theta_{kl}) = \log L_{kl}(\theta_{kl})$ is then

$$l_{kl}(\theta_{kl}) = \sum_{i=1}^{n} \int_0^{\infty} \bar{Y}_{ik}(u) \left\{ \log \lambda_{kl}(u|\mathcal{H}_i(u^-);\theta_{kl}) dN_{ikl}(u) - \lambda_{kl}(u|\mathcal{H}_i(u^-);\theta_{kl}) du \right\}, \quad (2.7)$$

where $\{N_{ikl}(u), u > 0\}$ is a counting process for $k \to l$ transitions for individual i.

For Markov models, recall that $\lambda_{kl}(t|\mathcal{H}_i(t^-);\theta_{kl}) = \lambda_{kl}(t;\theta_{kl})$. Now suppose you have a time-homogeneous model where $\lambda_{kl}(t;\theta_{kl}) = \theta_{kl}$, then

$$\begin{aligned} l_{kl}(\theta_{kl}) &= \sum_{i=1}^{n} \int_0^{\infty} \bar{Y}_{ik}(u) \left\{ \log \theta_{kl} dN_{ikl}(u) - \theta_{kl} du \right\} \\ &= \log \theta_{kl} \left(\sum_{i=1}^{n} \int_0^{\infty} \bar{Y}_{ik}(u) dN_{ikl}(u) \right) - \theta_{kl} \left(\sum_{i=1}^{n} \int_0^{\infty} \bar{Y}_{ik}(u) du \right) \\ &= \log \theta_{kl} n_{kl} - \theta_{kl} S_k \ , \end{aligned}$$

where n_{kl} is the total number of $k \to l$ transitions observed across individuals in the sample and $S_k = \sum_{i=1}^{n} \int_0^{\infty} \bar{Y}_{ik}(u) du$ is the total person-time at risk of transition out of state k. This partial log-likelihood has the same form as one for a time-homogeneous Poisson process where n_{kl} denotes the count and S_k denotes the exposure time.

The maximum likelihood estimate (MLE) for the $k \to l$ transition rate is $\hat{\theta}_{kl} = n_{kl}/S_k$. For time-homogeneous models, the estimated observed Fisher information matrix is $\hat{I}_{kl}(\hat{\theta}_{kl}) = n^{-1} n_{kl}/\hat{\theta}_{kl}^2$, so the estimated normal approximation for the distribution of $\hat{\theta}_{kl}$ is $\sqrt{n}(\hat{\theta}_{kl} - \theta_{kl}) \sim \mathrm{N}(0, n\hat{\theta}_{kl}^2/n_{kl})$. Also note that the likelihood ratio statistic is

$$\begin{aligned} LRS_{kl}(\theta_{kl}) &= 2\{\log L_{kl}(\hat{\theta}_{kl}) - \log L_{kl}(\theta_{kl})\} \\ &= 2\{n_{kl} \log(\hat{\theta}_{kl}/\theta_{kl}) - S_k(\hat{\theta}_{kl} - \theta_{kl})\}, \end{aligned}$$

where $LRS_{kl}(\theta_{kl}) \sim \chi^2_{(1)}$.

As the assumption of constant transition intensities is strong, we need more flexible functional forms for modelling intensities. A simple extension is a model with a piecewise constant intensity function. Let $0 = b_0 < b_1 < \cdots < b_R = \infty$ be a partition of the positive real line, with $\mathcal{B}_r = [b_{r-1}, b_r)$ and $\cup_{r=1}^{\infty}\mathcal{B}_r = [0, \infty)$. Then with a piecewise constant framework the intensities are $\lambda_{kl}(t; \theta_{kl}) = \theta_{klr},\ t \in \mathcal{B}_r,\ r = 1, \ldots, R$. The partial log-likelihood for parameters $\theta_{kl} = (\theta_{kl1}, \ldots, \theta_{klR})'$ is

$$\begin{aligned}\log L_{kl} &= \sum_{i=1}^{n}\sum_{r=1}^{R}\int_0^{\infty} \bar{Y}_{ikr}(u)\left\{\log \theta_{klr} dN_{ikl}(u) - \theta_{klr} du\right\} \\ &= \sum_{r=1}^{R}\left\{n_{klr}\log\theta_{klr} - S_{kr}\theta_{klr}\right\},\end{aligned}$$

where $\bar{Y}_{ikr}(u) = \bar{Y}_{ik}(u)I(u \in \mathcal{B}_r)$, $n_{klr} = \sum_{i=1}^{n}\int_0^{\infty}\bar{Y}_{ikr}(u)dN_{ikl}(u)$ is the total number of observed $k \to l$ transitions over $\mathcal{B}_r$, and $S_{kr} = \sum_{i=1}^{n}\int_0^{\infty}\bar{Y}_{ikr}(u)du$ is the total person-time at risk for $k \to l$ transitions. The MLE is $\hat{\theta}_{klr}$ and we have $\sqrt{n}(\hat{\theta}_{klr} - \theta_{klr}) \sim \mathrm{N}(0, n\hat{\theta}_{klr}^2/n_{klr})$.

Transition probabilities are functions of θ_{kl} and several approaches have been proposed to obtain variance estimates or confidence intervals for these. Suppose you are interested in $P_{kl}(s, t; \theta)$ where $\theta = (\theta_{kl},\ k = 1, \ldots, K,\ l = 1, \ldots, K)$ contains all elements of all the θ_{kl} parameters. By letting $\theta = (\theta_1, \ldots, \theta_m)'$ we are indexing all the θ_{kl} in some order where $m = K^2$ (i.e., the number of θ_{kl} parameters). An estimate for the asymptotic variance of the MLE is

$$\widehat{Var}(P_{kl}(s,t;\hat{\theta})) = \sum_{r=1}^{m}\sum_{u=1}^{m}\left\{\frac{\partial P_{kl}(s,t;\theta)}{\partial\theta_r}\frac{\partial P_{kl}(s,t;\theta)}{\partial\theta_u}\right\}\bigg|_{\hat{\theta}}\widehat{Cov}(\hat{\theta}_r, \hat{\theta}_u),$$

where $\widehat{Cov}(\hat{\theta}) = m^{-1}I^{-1}(\hat{\theta})$ is the estimated covariance matrix for $\hat{\theta}$ with elements $\widehat{Cov}(\hat{\theta}_r, \hat{\theta}_u)$ and the $m \times m$ matrix $I(\hat{\theta})$ is obtained from components of the estimated and normalized observed information matrix and the fact that separate $\hat{\theta}_{kl}$ are asymptotically independent.

For some Markov models there are simple expressions for P_{kl} where the derivatives can be determined analytically, however numerical approximations are generally required.

One effective approach is to use numerical differentiation where

$$\frac{\partial P_{kl}(s,t;\theta)}{\partial \theta_r} \doteq \frac{P_{kl}(s,t;\theta+\Delta_r) - P_{kl}(s,t;\theta-\Delta_r)}{2\delta_r},$$

where $\Delta_r = m \times 1$ zero vector except for the small value $\delta_r > 0$ for the element corresponding to θ_r. Note that one can also use the nonparametric bootstrap method for variance estimation (Cook and Lawless, 2018).

2.4.2 Nonparametric Estimation

For Markov models without predictors, the nonparametric estimation of cumulative transition intensities and other features of interest are possible.

Transition intensity functions take the form $\lambda_{kl}(t)$ for $k \neq l$, where $\lambda_{kk}(t) := -\sum_{l \neq k} \lambda_{kl}(t)$ for $k = 1, \ldots, K$. The nonparametric **Nelson-Aalen (NA) estimator** of the cumulative intensities $\Lambda_{kl}(t) = \int_0^t d\Lambda_{kl}(u) = \int_0^t \lambda_{kl}(u)du$ is

$$\begin{aligned}\hat{\Lambda}_{kl}(t) &= \sum_{i=1}^{n} \sum_{t_{ir} \in \mathcal{D}_{ikl}(t)} \frac{I(t_{ir} \leq t)}{\bar{Y}_{\cdot k}(t_{ir})} \\ &= \int_0^t d\hat{\Lambda}_{kl}(u) = \int_0^t \frac{d\bar{N}_{\cdot kl}(u)}{\bar{Y}_{\cdot k}(u)}, \ k \neq l,\end{aligned}$$

where $\bar{Y}_{\cdot k}(t) = \sum_{i=1}^{n} Y_{ik}(t)$. Thus $d\Lambda_{kl}(u)$ is estimated by the number of $k \to l$ transitions observed at time u divided by the number of individuals at risk for a transition out of state k (i.e., those under observation and in state k at time u^-). These estimators are analogous to the NA estimators of the cumulative hazard function in survival analysis and can be thought of as discrete MLEs where $\Lambda_{kl}(t)$ increases only at times where $k \to l$ transitions are observed.

We define $J_k(u) := I(\bar{Y}_{\cdot k}(u) > 0)$, where $J_k(u)/\bar{Y}_{\cdot k}(u)$ is defined to be 0 when $\bar{Y}_{\cdot k}(u) = 0$. For a continuous-time process with *no ties* (i.e., only one transition can simultaneously occur at a given time), the **variance estimate** is

$$\widehat{Var}(\hat{\Lambda}_{kl}(t)) = \int_0^t \frac{J_k(u)d\hat{\Lambda}_{kl}(u)}{\bar{Y}_{\cdot k}(u)} = \sum_{t^{(r)} \le t} \frac{J_k(t^{(r)})}{\bar{Y}_{\cdot k}(t^{(r)})^2} d\bar{N}_{\cdot kl}(t^{(r)}),$$

where $t^{(1)} < \cdots < t^{(m)}$ are distinct times at which m observed transitions occur. However, in practice occasional ties can occur because the transition times are recorded on a discrete time scale. An alternative estimate based on a discrete time framework for handling *ties* is

$$\begin{aligned}\widehat{Var}(\hat{\Lambda}_{kl}(t)) &= \sum_{t^{(r)} \le t} \frac{J_k(t^{(r)})d\bar{N}_{\cdot kl}(t^{(r)}) \left(\bar{Y}_{\cdot k}(t^{(r)}) - d\bar{N}_{\cdot kl}(t^{(r)})\right)}{\bar{Y}_{\cdot k}(t^{(r)})^3} \\ &= \sum_{t^{(r)} \le t} \frac{J_k(t^{(r)})d\hat{\Lambda}_{kl}(t^{(r)}) \left(1 - d\hat{\Lambda}_{kl}(t^{(r)})\right)}{\bar{Y}_{\cdot k}(t^{(r)})}.\end{aligned}$$

These two variance estimates are close in value when large numbers $\bar{Y}_{\cdot k}(t^{(r)})$ are at risk and there are few ties. We also assume that as $n \to \infty$, $J_k(u) > 0$ with probability one.

The nonparametric estimates of the matrix of transition probabilities are

$$\hat{P}(s,t) = \prod_{(s,t]} \{I + \hat{Q}(u)du\}, \ (= \textbf{Aalen-Johansen estimator})$$

where $\hat{Q}(u)du$ is a $K \times K$ matrix with off-diagonal entries $d\hat{\Lambda}_{kl}(u)$ given by the integrand from the NA estimator and diagonal entries $-\sum_{l \neq k} d\hat{\Lambda}_{kl}(u), \ k = 1, \ldots, K$. Note that the R package `etm` (Allignol et al., 2011) provides these estimates and their variance estimates.

2.5 Mixed Effects Models

A mixed effects model contains both fixed effects and random effects. The *linear mixed-effects models* have the form

$$\boldsymbol{y} = X\boldsymbol{\beta} + Z\boldsymbol{b} + \boldsymbol{\epsilon}, \quad \boldsymbol{b} \sim \mathrm{N}(\boldsymbol{0}, \psi_{\boldsymbol{\theta}}), \quad \boldsymbol{\epsilon} \sim \mathrm{N}(\boldsymbol{0}, \sigma^2\Lambda), \tag{2.8}$$

where $\boldsymbol{y} = (y_1, \ldots, y_n)'$ is an $n \times 1$ column vector of observations on the vector of random variables $\boldsymbol{Y} = (Y_1, \ldots, Y_n)'$, X is the model matrix for the fixed effects, $\boldsymbol{\beta}$ is the vector containing the coefficients associated with the fixed effects, Z is the model matrix for the random effects, $\boldsymbol{b}$ is the random vector containing *random effects* with mean zero and positive definite covariance matrix $\psi_{\boldsymbol{\theta}}$ (unknown parameters $\boldsymbol{\theta}$), $\boldsymbol{\epsilon} = (\epsilon_1, \ldots, \epsilon_n)'$ is the $n \times 1$ column vector containing the error terms that are mutually independent zero mean random variables with the same variance σ^2, and Λ is a positive definite matrix, of simple structure, which is typically used to model residual autocorrelation. Often, Λ is the identity matrix which corresponds to the assumption that the random effects are uncorrelated.

Linear mixed-effects models allow for a more complex stochastic structure compared to fixed-effects models, and imply that the elements of the response are no longer independent. The model form from (2.8) can be rewritten as

$$\boldsymbol{y} = X\boldsymbol{\beta} + \boldsymbol{e}, \quad \boldsymbol{e} = Z\boldsymbol{b} + \boldsymbol{\epsilon} \sim \mathrm{N}(\boldsymbol{0}, \sigma^2\Sigma_{\boldsymbol{\theta}}), \tag{2.9}$$

where $\Sigma_{\boldsymbol{\theta}} = \frac{Z\psi_{\boldsymbol{\theta}}Z'}{\sigma^2} + I$ since the covariance of $\boldsymbol{e}$ is equal to $Z\psi_{\boldsymbol{\theta}}Z' + \sigma^2 I$ and I is the identity matrix. The likelihood for this linear mixed-effects model is

$$\mathcal{L}(\boldsymbol{\beta}, \boldsymbol{\theta}, \sigma^2) = \frac{1}{\sqrt{(2\pi\sigma^2)^n \mid \Sigma_{\boldsymbol{\theta}} \mid}} \exp\left[-(\boldsymbol{y} - X\boldsymbol{\beta})'\Sigma_{\boldsymbol{\theta}}^{-1}(\boldsymbol{y} - X\boldsymbol{\beta})/(2\sigma^2)\right], \tag{2.10}$$

where maximizing $\mathcal{L}$ with respect to $\boldsymbol{\beta}, \boldsymbol{\theta}, \sigma^2$ will provide $\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\theta}}, \hat{\sigma}^2$. However, this maximization can be simplified by profiling the likelihood (not shown here).

The restricted maximum likelihood (REML) approach is an estimation technique that takes the average of the joint likelihood $\mathcal{L}(\boldsymbol{\beta}, \boldsymbol{\theta}, \sigma^2)$ over all possible values of $\boldsymbol{\beta}$, called the REML criterion, and is maximized to find the variance parameters $\hat{\boldsymbol{\theta}}, \hat{\sigma}^2$. The REML criterion is

$$\begin{aligned}
\mathcal{L}_R(\boldsymbol{\theta}, \sigma^2) &= \int \mathcal{L}(\boldsymbol{\beta}, \boldsymbol{\theta}, \sigma^2) d\boldsymbol{\beta} \\
&= \frac{1}{\sqrt{(2\pi)^n \mid \Sigma \mid}} \int \exp\left[-(\boldsymbol{y} - X\boldsymbol{\beta})' \Sigma^{-1} (\boldsymbol{y} - X\boldsymbol{\beta})/2\right] d\boldsymbol{\beta}, \qquad \Sigma = \Sigma_{\boldsymbol{\theta}} \sigma^2 \\
&= \dots \\
&= \frac{\exp\left[-(\boldsymbol{y} - X\hat{\boldsymbol{\beta}})' \Sigma^{-1} (\boldsymbol{y} - X\hat{\boldsymbol{\beta}})/2\right]}{\sqrt{(2\pi)^n \mid \Sigma \mid}} \sqrt{\frac{(2\pi)^p}{\mid X' \Sigma^{-1} X \mid}},
\end{aligned}$$

where p is the dimension of $\boldsymbol{\beta}$ and n is the dimension of $\boldsymbol{y}$. For more details see Wood (2006).

2.6 Frailty Models

Frailty models are random effects hazard models where the random effect (i.e., frailty) has a multiplicative effect on the hazard. The hazard can depend on predictors which can be modeled in a parametric or semi-parametric way (Duchateau and Janssen, 2007). Frailty models extend the Cox PH model to account for unobservable heterogeneity among individuals.

For subject j, $j = 1, \dots, n_i$ from cluster[3] i, $i = 1, \dots, s$, let the observation y_{ij} be the minimum of the censoring time and the event time. The shared frailty model has the

[3] A **cluster** is a group of objects that are more similar to each other than to those in other groups, or clusters.

form

$$h_{ij}(t) = h_0(t)\exp(\boldsymbol{x}_{ij}^{'}\boldsymbol{\beta} + b_i)$$
$$= h_0(t)u_i \exp(\boldsymbol{x}_{ij}^{'}\boldsymbol{\beta}),$$

where $h_{ij}(t)$ is the conditional hazard function (conditional on b_i), $h_0(t)$ is the baseline hazard function, $\boldsymbol{x}_{ij}^{'}$ is the $1 \times p$ row vector of predictors for subject j from cluster i, $\boldsymbol{\beta}$ is the $p \times 1$ column vector containing the coefficients associated with the fixed effects (also called the fixed effects vector), and b_i is the random effect for the ith cluster (Duchateau et al. (2002), Rondeau (2010)). The frailty term, $u_i = \exp(b_i)$, has a multiplicative effect on the baseline hazard function.

2.6.1 Frailty Distributions

Duchateau et al. (2002) provide two common choices for the density of frailties $u = \exp(b)$:

1. The zero-mean Normal density for b (i.e., the **Lognormal** density for u), where

$$f_U(u) = \frac{1}{u\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(\log(u))^2}{2\sigma^2}\right), \qquad \mu = 0,\ \sigma > 0.$$

2. The one-parameter **Gamma** density for u, where

$$f_U(u) = \frac{u^{(1/\theta)-1}\exp(-u/\theta)}{\theta^{1/\theta}\Gamma(1/\theta)}, \qquad \theta > 0.$$

For the one-parameter Gamma density function the scale parameter is θ and the shape parameter is $1/\theta$ resulting in $E(u) = 1$ and $Var(u) = \theta$. The Gamma distribution is the most commonly used finite mean distribution to model the frailty term in such models (Ibrahim et al., 2001, p. 101). The parameter θ gives information on the variability (i.e., the heterogeneity) in the population of clusters — larger values of θ imply greater heterogeneity among clusters.

For frailty models, the baseline hazard can be specified using the functions discussed in Chapter 2.2.1. See Ibrahim et al. (2001) and Austin (2017) for further details.

2.7 Bayesian Methods

Bayesian data analysis methods assume that the parameters for parametric models are random variables as opposed to the fixed (but unknown) constants that are used in frequentist modelling frameworks. These methods use probability for quantifying uncertainty in inferences and hypotheses and the probabilities are updated using Bayes' rule as more information becomes available. In general, the method follows these three steps:

1. Setting up a joint probability distribution for all quantities where any previous knowledge of the problem can be utilized,

2. Calculating the conditional probability distribution of those unobserved quantities of interest given the data that is observed, and

3. Evaluating the fit of the model to determine if the conclusions are reasonable and consistent with the data.

One strength of these methods is the interpretability of the results since Bayesian (or probability) intervals[4] for unknown quantities are interpreted as having a high probability of containing said unknown quantities whereas frequentist (or confidence) intervals for unknown quantities are interpreted as a range of values that contains said unknown quantities, some percentage (e.g., 95%) of the time.

We provide a brief overview of Bayesian methods. For more information about Bayesian data analysis see Gelman et al. (2014a) and Kruschke (2015).

[4] Also known as credible intervals.

Suppose we have some observable data D and we want to make inferences about some parameter θ. We use the notation $p(\cdot)$ to denote a marginal probability density and $p(\cdot|\cdot)$ to denote a conditional probability density, where the arguments in both densities are determined by the problem.

The *joint probability density function* can be written in the form

$$p(\theta, D) = p(\theta)p(D|\theta) = p(\theta|D)p(D), \tag{2.11}$$

where $p(\theta)$ is the prior distribution (explained in the next section) and $p(D|\theta)$ is the likelihood of the data conditional on the parameter. Using Bayes' rule, we obtain the *posterior* density

$$p(\theta|D) = \frac{p(\theta, D)}{p(D)} = \frac{p(\theta)p(D|\theta)}{p(D)}, \tag{2.12}$$

where $p(D) = \sum_\theta p(\theta)p(D|\theta)$, which is a sum over all possible values of θ, for discrete θ or $p(D) = \int p(\theta)p(D|\theta)d\theta$ for continuous θ. Since $p(D)$ in the denominator of (2.12) does not depend on θ and can be viewed as a constant for fixed D, we can simplify the right-hand side of (2.12) to

$$\begin{aligned} p(\theta|D) &\propto p(\theta)p(D|\theta) \\ &= \text{prior distribution} \times \text{likelihood function.} \end{aligned}$$

We see that the posterior density is proportional to the prior distribution times the likelihood function; we use our prior knowledge or beliefs and the likelihood of the observed data to determine our posterior distribution (or density or probability).

2.7.1 Priors and Hyperpriors

A prior probability distribution, or simply a prior, denoted above by $p(\theta)$ allows us to use some knowledge or belief that we already have about the unknown quantity θ before any evidence is taken into account. For example, if we wanted to find the probability of selling lemonade on a hot and sunny day, we can use our prior knowledge about the likelihood of selling lemonade on any type of day (e.g., rainy, windy, snowy, etc.).

There are many types of priors to consider and we will only mention a few of them here. A *non-informative prior*[5] is one that represents complete ignorance about the value of the parameter with as few restrictions on the parameter space as possible. The rationale for using these types of priors is to allow the data to speak for itself. A *weakly informative prior* is one which contains some information about a parameter but still allows the likelihood to dominate the posterior. An *informative prior* expresses specific and/or substantive information about a parameter. Such priors incorporate expert information, reduce the variance of the posterior, and improve simulation-based estimation (discussed below), but misspecified informative priors or priors with too small of a variance can have negative effects on the results. See Chapter 2 in Gelman et al. (2014a) for more details about prior distributions, along with other types (e.g., conjugate and improper priors).

Prior distibutions can have parameters in them, and these parameters are called *hyperparameters*. For instance, suppose we want to toss a coin one time. The probability of the coin turning up heads is denoted by p. Before we toss the coin, you believe that the coin is fair (i.e., $p = 0.5$) rather than biased and I believe that all values of $p \in [0, 1]$ are equally plausible. Our prior beliefs about p can be modeled using a Beta distribution where $p \sim \mathrm{B}(\alpha, \beta)$. Thus, the α and β parameters (or the shape and scale parameters, respectively) of the prior Beta distribution are called hyperparameters. These hyperparameters may also have distributions that express prior beliefs about their values. The prior distributions of hyperparameters are known as *hyperpriors*.

[5]Non-informative priors may make assumptions about the structure of the parameter space, such as assuming independence between parameters or limiting parameters to a finite value. A prior may be non-informative on one scale but informative after transformation. For example, a uniform distribution on a variance represents ignorance and might appear non-informative. However, it is informative for the inverse variance, also known as the precision.

2.7.2 Gibbs Sampling

Gibbs sampling, also called the alternating conditional sampling, is a Markov chain Monte Carlo (MCMC) algorithm that is used when the joint distribution (2.11) either has no closed form or is too difficult to calculate directly. This also occurs for the marginal likelihood shown in the denominator in (2.12) since it is rarely available in closed form. Instead, MCMC methods like the Gibbs sampler are employed to sample from the known conditional posterior distribution, $p(\theta|D)$, for each parameter and these distributions are often much easier to sample from. Thus, the Gibbs sampler allows one to obtain a sequence of observations which are estimated from conditional probability distributions without requiring the difficult derivation of and estimation from the joint probability distribution.

Gibbs sampling is a special case of the Metropolis-Hastings (MH) family of samplers. The Metropolis-Hastings algorithm is an adaptation of a random walk with an acceptance/rejection rule to draw samples from the posterior. We provide a brief overview of one version of this algorithm. Suppose we have data $\boldsymbol{y} = (y_1, \ldots, y_n)'$, a vector of random variables $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_k)'$, and a distribution function of the data $p_Y(\boldsymbol{y}|\boldsymbol{\theta})$. We want to obtain a sample from the joint posterior distribution, $p_\Theta(\boldsymbol{\theta}|\boldsymbol{y})$. From (2.12) we have

$$p_\Theta(\boldsymbol{\theta}|\boldsymbol{y}) = \frac{p_\Theta(\boldsymbol{\theta})p_Y(\boldsymbol{y}|\boldsymbol{\theta})}{\int_\Theta p_\Theta(\boldsymbol{\theta})p_Y(\boldsymbol{y}|\boldsymbol{\theta})d\boldsymbol{\theta}} \propto p_\Theta(\boldsymbol{\theta})p_Y(\boldsymbol{y}|\boldsymbol{\theta}).$$

An intuitive approach to the steps of the algorithm are:

1. Initialize: Start with a vector of initial values where $\boldsymbol{\theta}^{(0)} = (\theta_1^{(0)}, \ldots, \theta_k^{(0)})'$.

2. Propose: For each initial value $\theta_i^{(0)}$ propose a new value θ_i^* based on a small deviation from the current value. In the first iteration, the proposed values will be based on the initial values. The vector of proposals is denoted by $\boldsymbol{\theta}^*$.

3. Calculate: Using the proposed and current values, calculate the acceptance ratio

$$r = p_\Theta(\boldsymbol{\theta}^*|\boldsymbol{y})/p_\Theta(\boldsymbol{\theta}^{(0)}|\boldsymbol{y}) = \frac{p_\Theta(\boldsymbol{\theta}^*)p_Y(\boldsymbol{y}|\boldsymbol{\theta}^*)}{p_\Theta(\boldsymbol{\theta}^{(0)})p_Y(\boldsymbol{y}|\boldsymbol{\theta}^{(0)})}.$$

4. Decide: If $r > 1$ then $\boldsymbol{\theta}^*$ makes the data more likely than $\boldsymbol{\theta}^{(0)}$. We accept the proposal and set $\boldsymbol{\theta}^{(1)} = \boldsymbol{\theta}^*$. If $r < 1$ then we set $\boldsymbol{\theta}^{(1)}$ to either $\boldsymbol{\theta}^*$ or $\boldsymbol{\theta}^{(0)}$ with probability r or $1 - r$, respectively.

This process is repeated using steps 2-4 where the steps are updated to the following:

2. Propose: For each value $\theta_i^{(repetition)}$, where $repetition = 1, 2, \ldots$, propose a new value θ_i^* based on the current value. The vector of proposals is denoted by $\boldsymbol{\theta}^*$.

3. Calculate: Using the proposed and current values, calculate the acceptance ratio

$$r = p_\Theta(\boldsymbol{\theta}^*|\boldsymbol{y})/p_\Theta(\boldsymbol{\theta}^{(repetition)}|\boldsymbol{y}) = \frac{p_\Theta(\boldsymbol{\theta}^*)p_Y(\boldsymbol{y}|\boldsymbol{\theta}^*)}{p_\Theta(\boldsymbol{\theta}^{(repetition)})p_Y(\boldsymbol{y}|\boldsymbol{\theta}^{(repetition)})}.$$

4. Decide: If $r > 1$ then $\boldsymbol{\theta}^*$ makes the data more likely than $\boldsymbol{\theta}^{(repetition)}$. We accept the proposal and set $\boldsymbol{\theta}^{(repetition+1)} = \boldsymbol{\theta}^*$. If $r < 1$ then we set $\boldsymbol{\theta}^{(repetition+1)}$ to either $\boldsymbol{\theta}^*$ or $\boldsymbol{\theta}^{(repetition)}$ with probability r or $1 - r$, respectively.

The repetition continues until a desired number of samples is reached. Note that the algorithm tends to stay in and return large numbers of samples from high-density regions of the posterior distribution while only occasionally visiting low-density regions. An MCMC routine is said to have converged when it is sampling from the highest density regions, rather than exploring the parameter space for these regions. One important disadvantage of this algorithm to highlight is that it may take a long time for the Markov chain to converge if the initial values are in a region of low density.

Recall that Gibbs sampling is used when the conditional posterior distribution is known and easier to sample from than the joint posterior distribution. Thus, the

Metropolis-Hastings algorithm samples from the joint posterior distribution whereas the Gibbs algorithm samples from the conditional posterior distribution. For Gibbs sampling, we set initial values for $\boldsymbol{\theta}_{-1} = (\theta_2, \ldots, \theta_k)'$, i.e., for all but the first parameter. Then there are k steps in each iteration where the sampler cycles through the subvectors of $\boldsymbol{\theta}$ and draws each subset conditional on the value of all the others. The procedure of one iteration follows:

1. Sample θ_1^* from $p_{\theta_1}(\theta_1|\theta_2^{(0)}, \theta_3^{(0)}, \ldots, \theta_{j-1}^{(0)}, \theta_j^{(0)}, \theta_{j+1}^{(0)}, \ldots, \theta_k^{(0)}, \boldsymbol{y})$. Accept the proposal using the Metropolis-Hastings rules and set the value of $\theta_1^{(1)}$.

2. Sample θ_2^* from $p_{\theta_2}(\theta_2|\theta_1^{(1)}, \theta_3^{(0)}, \ldots, \theta_{j-1}^{(0)}, \theta_j^{(0)}, \theta_{j+1}^{(0)}, \ldots, \theta_k^{(0)}, \boldsymbol{y})$. Again, accept the proposal using the rules and set $\theta_2^{(1)}$.

3. Continue sampling θ_j^*, $j = 3, \ldots, k$, from $p_{\theta_j}(\theta_j|\theta_1^{(1)}, \theta_2^{(1)}, \theta_3^{(1)}, \ldots, \theta_{j-1}^{(1)}, \theta_{j+1}^{(0)}, \ldots, \theta_k^{(0)}, \boldsymbol{y})$. until a sample of $\boldsymbol{\theta}^{(1)} = (\theta_1^{(1)}, \theta_2^{(1)}, \ldots, \theta_k^{(1)})'$ is found.

Repeat this process by moving to the next iteration, starting with the value of $\theta_1^{(2)}$. Stop at the desired number of samples. If the model converges then the samples from the Gibbs sampler approximate the joint distribution of all the parameters. Thus, the sample properties, such as the mean, median, mode, variance, and covariance between parameters, from the Gibbs sampler are the same as the sample properties of samples from the joint distribution. For more details on these samplers or MCMC simulation in general, see Chapter 11 in Gelman et al. (2014a) and Chapter 7 in Kruschke (2015).

JAGS (Just Another Gibbs Sampler) is a program for Bayesian modelling that builds MCMC samplers for complex hierarchical models (Plummer, 2003). It succeeded the BUGS (Bayesian inference Using Gibbs Sampling) system and has retained many of the features of BUGS but utilizes different samplers and has better useability across various operating systems like Linux, Windows and MacOS. JAGS inputs a description of a model for the data and outputs an MCMC sample of the posterior distribution. Several packages exist to connect the statistical software R (R Core Team, 2021) with JAGS like the `runjags` package (Denwood, 2016).

2.8 Evaluation of Bayesian Models

After a Bayesian model is fit using the methods mentioned above, we often want to measure its predictive accuracy especially when performing model comparisons. Suppose the data $\boldsymbol{y} = (y_1, \ldots, y_n)'$ are independent given parameters $\boldsymbol{\theta}$. Then the likelihood function is $p(\boldsymbol{y}|\boldsymbol{\theta}) = \sum_{i=1}^{n} p(y_i|\boldsymbol{\theta})$, the posterior distribution is $p(\boldsymbol{\theta}|\boldsymbol{y})$, the prior distribution is $p(\boldsymbol{\theta})$, and the *posterior predictive density* (or *distribution*) is $p(\widetilde{\boldsymbol{y}}|\boldsymbol{y}) = \int p(\widetilde{y}_i|\boldsymbol{\theta})p(\boldsymbol{\theta}|\boldsymbol{y})d\boldsymbol{\theta}$ since it is a prediction for an observable $\widetilde{\boldsymbol{y}}$. The expected log pointwise predictive density (ELPD) for a new dataset provides a measure of predictive accuracy for the n data points taken one at a time, defined by

$$\text{ELPD} = \sum_{i=1}^{n} \int p_t(\widetilde{y}_i) \log p(\widetilde{y}_i|\boldsymbol{y}) d\widetilde{y}_i, \tag{2.13}$$

where $p_t(\widetilde{y}_i)$ is the unknown distribution of the true data-generating process for $\widetilde{y}_i$. The methods described below approximate this equation.

Note that the definitions stated in this chapter are those used in the `loo` package (Vehtari et al., 2020) and discussed in the paper by Vehtari et al. (2017).

2.8.1 Widely Applicable Information Criterion

The Widely Applicable Information Criterion (WAIC), also referred to as the Watanabe-Akaike Information Criterion (Watanabe and Opper, 2010), is one of the methods used to compare Bayesian models. It it different than the commonly used Deviance Information Criterion (DIC) since it is fully Bayesian. The DIC (Gelman et al., 2014b) is based on a point estimate whereas the WAIC uses the entire posterior distribution, making it a fully Bayesian criterion and one that is asymptotically equivalent to Bayesian cross-validation.

It is an alternative approach to estimating the expected log pointwise predictive density shown in (2.13).

WAIC is composed of two parts: the log pointwise predictive density and the effective number of parameters. The log pointwise predictive density (LPD), an overestimate[6] of (2.13), is defined as

$$\begin{aligned}\mathrm{LPD} &= \sum_{i=1}^{n} \log p_(y_i|\boldsymbol{y}) = \sum_{i=1}^{n} \log \int p(y_i|\boldsymbol{\theta})p(\boldsymbol{\theta}|\boldsymbol{y})d\boldsymbol{\theta} \\ \widehat{\mathrm{LPD}} &= \sum_{i=1}^{n} \log \left(\frac{1}{S} \sum_{s=1}^{S} p(y_i|\boldsymbol{\theta}^s) \right)\end{aligned}$$

where $\boldsymbol{\theta}^s$ is the sth draw out of S total draws from the posterior distribution based on an MCMC algorithm. The effective number of parameters (p_{WAIC}) is a penalty for having too many unconstrained parameters and is defined as

$$\begin{aligned}p_{\mathrm{WAIC}} &= \sum_{i=1}^{n} \mathrm{var}_{\mathrm{post}} \left(\log p(y_i|\boldsymbol{\theta})\right) \\ \widehat{p}_{\mathrm{WAIC}} &= \sum_{i=1}^{n} V_{s=1}^{S} \left(\log p(y_i|\boldsymbol{\theta}^s)\right)\end{aligned}$$

where $V_{s=1}^{S}(a_s) = \frac{\sum_{s=1}^{S}(a_s - \bar{a})^2}{S-1}$ is the sample variance function. Note that the definition of the effective number of parameters corresponds to p_{WAIC2} in Gelman et al. (2014b). The theoretical and computing formulas for the WAIC are

$$\begin{aligned}\mathrm{WAIC} &= -2 \times \mathrm{ELPD} = -2(\mathrm{LPD} - p_{\mathrm{WAIC}}) \qquad (2.14) \\ \widehat{\mathrm{WAIC}} &= -2 \times \widehat{\mathrm{ELPD}} = -2(\widehat{\mathrm{LPD}} - \widehat{p}_{\mathrm{WAIC}}).\end{aligned}$$

We can see that the WAIC is more stable than the DIC from its definition since it computes the sample variance of the draws separately for each data point and then sums

[6] An estimate is an **overestimate** if it exceeds the actual result. In contrast, an estimate is an underestimate if it is less than the actual result.

over them.

2.8.2 Pareto Smoothed Importance Sampling

The Bayesian leave-one-out (LOO) estimate of the out-of-sample predictive fit (i.e., the leave out one data point estimate for 2.13) is

$$\text{ELPD}_{\text{LOO}} = \sum_{i=1}^{n} \log p(y_i|\boldsymbol{y}_{-i}),$$

where $p(y_i|\boldsymbol{y}_{-i}) = \int p(y_i|\boldsymbol{\theta})p(\boldsymbol{\theta}|\boldsymbol{y}_{-i})d\boldsymbol{\theta}$ is the leave-one-out predictive density given the data without the ith data point. This estimate can be greatly improved upon using Pareto smoothed importance sampling (PSIS) which we discuss below.

Raw importance sampling evaluates $p(y_i|\boldsymbol{y}_{-i})$ using importance ratios, r_i^s, which are defined as

$$r_i^s = \frac{1}{p(y_i|\boldsymbol{\theta}^s)} \propto \frac{p(\boldsymbol{\theta}^s|\boldsymbol{y}_{-i})}{p(\boldsymbol{\theta}^s|\boldsymbol{y})},$$

where $\boldsymbol{\theta}^s$ is the sth draw out of S total draws from the posterior distribution (i.e., s indexes the simulation draws). Then the importance sampling leave-one-out (IS-LOO) predictive distribution is

$$p(\tilde{y}_i|\boldsymbol{y}_{-i}) \approx \frac{\sum_{s=1}^{S} r_i^s p(\tilde{y}_i|\boldsymbol{\theta}^s)}{\sum_{s=1}^{S} r_i^s},$$

and evaluating it at y_i, the held out data point, leads to

$$p(y_i|\boldsymbol{y}_{-i}) \approx \frac{1}{\frac{1}{S}\sum_{s=1}^{S}(p(y_i|\boldsymbol{\theta}^s))^{-1}}.$$

However, it is unstable because the importance ratios can have high or infinite variance. To deal with this issue we turn to truncated importance sampling which modifies the importance ratios. The $\boldsymbol{r}^s$ ratios are replaced by truncated weights $\boldsymbol{w}^s = \min(\boldsymbol{r}^s, S^{-1/2}\sum_{s=1}^{S}\boldsymbol{r}^s)$, which leads to having finite variance but also introduces bias.

Thus, we turn to Pareto smoothed importance sampling. A brief overview of the procedure follows but for more information see Section 2.1 in Vehtari et al. (2017) and Vehtari et al. (2021). The steps are:

1. Fit a generalized Pareto distribution to the 20% largest importance ratios, $r_i^s,\ s = 1, \ldots, S$, from the raw importance sampling. Do this separately for each held-out data point i.

2. Create new weights, $\widetilde{w}_i^s,\ s = 1, \ldots, S$, by retaining the $M - 1$ smallest ratios and replacing the M largest ratios with the expected values of the order statistics of the fitted generalized Pareto distribution, using the inverse-CDF: $F^{-1}\left(\frac{z-1/2}{M}\right),\ z = 1, \ldots, M$, where M is the number of simulation draws used to fit the Pareto (i.e., $M = 0.2S$).

3. Truncate each vector of weights at $S^{3/4}\overline{w}_i$, where $\overline{w}_i$ is the average of the S smoothed weights corresponding to the distribution without data point i. Label the weights as w_i^s.

These steps are done for each data point i and result in a vector of weights, $\boldsymbol{w}^s,\ s = 1, \ldots, S$, with elements $w_i^s,\ i = 1, \ldots, n$, that should behave better than the original importance ratios. Therefore the Pareto smoothed importance sampling estimate of the leave-one-out expected log pointwise predictive density is

$$\widehat{\text{ELPD}}_{\text{PSIS-LOO}} = \sum_{i=1}^{n} \log\left(\frac{\sum_{s=1}^{S} w_i^s p(y_i|\boldsymbol{\theta}^s)}{\sum_{s=1}^{S} w_i^s}\right). \tag{2.15}$$

The estimated shape parameter $\hat{k}$ of the generalized Pareto distribution is used to assess the reliability of the estimate in (2.15). If

$$\begin{cases} k < 1/2 & \implies \text{estimate converges quickly} \\ 1/2 \leq k \leq 1 & \implies \text{estimate converges slower} \\ k > 1 & \implies \text{more issues} \end{cases}.$$

Ultimately, if any $\hat{k} > 0.7$ then one should: (1) sample directly from $p(\boldsymbol{\theta}^s|\boldsymbol{y}_{-i})$ so long as the number of problematic data points is not too high; (2) use k-fold cross-validation; or (3) use a more robust model.

2.9 Knowledge Exchange

Terms like knowledge transfer and knowledge translation are often used when discussing the sharing or dissemination of knowledge, and are gaining prominence in Canada (Graham et al., 2006). Knowledge transfer or translation of research is the process in which knowledge is passed from researchers to stakeholders (e.g., consumers, end users, decision makers, etc.) or from stakeholders to researchers. Some people using these terms consider the passing of knowledge or information to be a two-way process but this is not always clear or explicit. Also, the idea of transferring knowledge can be limiting because it fails to encompass the goal of putting such knowledge into action.

The University of Western Ontario's Research Services assists researchers by providing financial support, ensuring oversight and compliance, and promoting the importance of research, to name a few. They maintain a webpage on "Knowledge Exchange and Impact" `https://www.uwo.ca/research/services/kex/index.html` which states that they support knowledge mobilization, translation and dissemination activities across the university. The webpage provides a brief overview about knowledge exchange and its benefits, without any clear definition. It states that heightened engagement between all partners and greater connections and collaborations are some of the benefits but does not emphasize the importance of engaging with all partners from the initial problem identification stage or throughout any subsequent stages of research. The webpage directs readers to additional resources on knowledge translation and knowledge impact (Grimshaw et al., 2012; Kothari and Wathen, 2017) but not on knowledge exchange.

Ward et al. (2012) developed a conceptual framework of knowledge exchange by embedding a knowledge broker[7] within a mental health organization in the UK where three teams were tackling various organizational tasks. Observational fieldnotes and interview data were collected and analyzed to refine their five components of knowledge exchange: **problem** identification and communication, analysis of **context**, **knowledge** development and selection, knowledge exchange activities or **interventions**, and knowledge **use**. These components do not necessarily have to occur in any particular order and multiple components may be occurring at the same time.

Alternatively, McFayden et al. (2022) [submitted] defined knowledge exchange as: (1) the collective overarching process where knowledge is collaboratively created, shared, and transformed as it is shared; and (2) the context in which people learn about knowledge. They conceptualize the process of knowledge exchange with a clear illustration of all the components (e.g., researchers, practitioners, and knowledge brokers), outlining the complex and intertwining relationships that exist. This process is further contextualized by their explanation of knowledge exchange within the field of fire management. They highlight certain aspects of this process such as research and development which relates directly to our work. It is clear that the research and development process within fire management both requires and benefits from knowledge exchange. This two-way, collaborative communication is present from the initial brainstorming stage until the application, or potential implementation, of findings stage.

2.10 Active Learning

Garfield (1995) proposed ten general principles of how students learn statistics. Such principles involve students learning by constructing their own knowledge instead of passively absorbing it (i.e., knowledge exchange instead of knowledge transfer), active involvement in learning activities, practicing concepts multiple times in different contexts, and becom-

[7] A **knowledge broker** is an individual or organization that provides a link between researchers and end users by creating a mutual understanding of goals. They help to facilitate the research process, from identifying a problem to translating the findings into policy and practice (Ward et al., 2012).

ing aware of and confronting their misconceptions, to name a few. These principles were regrouped into eight evidence-based recommendations for how students learn statistics in Garfield and Ben-Zvi (2007). Activity-based courses and active learning activities help students to learn statistics, but what does active learning truly mean?

Bonwell and Eison (1991) defined active learning strategies as "instructional activities involving students in *doing* things and *thinking* about what they are doing" (p. 7). These strategies allow students to engage with course material on a deeper level by becoming active, hands-on learners rather than passive ones. Figure 2.1 illustrates this definition for active learning along with stating some of the benefits. It is crucial to remember that active learning strategies are not a replacement of traditional lectures but should be used in conjunction with them — incorporated into lecture time or occasionally used in place of a lecture (e.g., a flipped classroom[8]) as required (Zakrajsek, 2018; French and Kennedy, 2017).

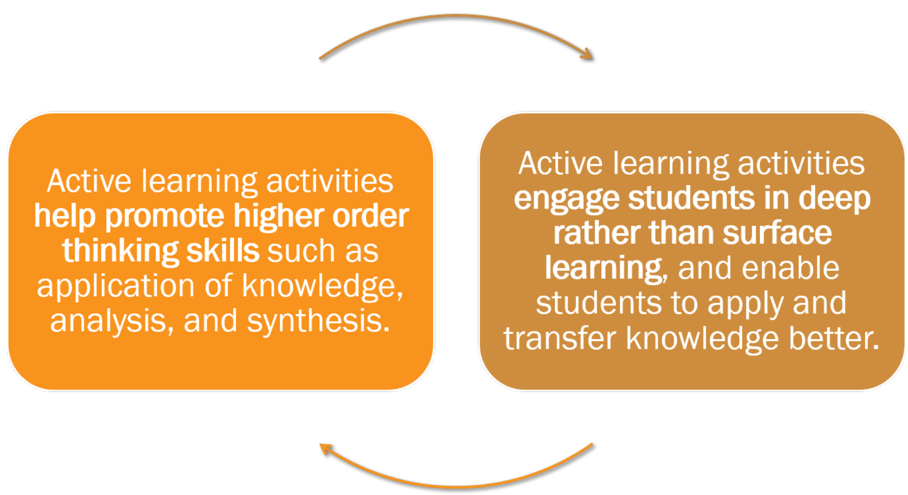


Figure 2.1: Definition of Active Learning. Taken from `https://www.queensu.ca/teachingandlearning/modules/active/04_what_is_active_learning.html`

[8]A **flipped classroom** is one where students review materials (e.g., videos, readings, online modules, etc.) before class time. Usually active learning activities are used during class time to allow students the opportunity to apply their new knowledge and refine their understanding of the material. The flipped classroom can be implemented during a single class or throughout an entire course.

Years of research and investigation into teaching and learning in higher education supports that active learning enhances student learning in many ways. The American Statistical Association's Guidelines for Assessment and Instruction in Statistics Education (GAISE) College Report (Carver et al., 2016) endorsed the use of active learning within the statistical community by recommending that statistics courses should foster active learning. Active learning strategies help students retain knowledge better (i.e., they perceive themselves as gaining more knowledge and understanding from their courses), prioritizes their growth and development within the classroom, and highlights the commitment that their university or college has to their welfare (Braxton et al., 2008). Students who engaged in active learning activities reported personal development like tolerance for obstacles and having an increased motivation to learn (Lopatto, 2007), and academic development with enhanced test scores and consistently higher test scores when compared with traditional learners (Mello and Less, 2013).

Active learning is often used as an umbrella term that encompasses other types of learning like collaborative learning (students working in groups to reach a common goal where the emphasis is placed on their interactions), cooperative learning (students working in groups but being assessed individually), and problem-based learning (occurs when a problem is introduced to provide motivation for the learning that follows). Evidence for the effectiveness of these types of learning styles are supported by Prince (2004) who notes that the improved learning is due to the nature of active engagement rather than the extra time spent on a specific topic. Importantly, in an inherently interdisciplinary discipline such as statistics, active learning, with its emphasis on collaboration, lays the foundation of the importance of collaboration in learning with spill-over effects that are important to collaboration in research.

Active learning fits perfectly into Dale's *Cone of Experience* (Dale, 1969) where people remember:

- 10% of what they **read**

- 20% of what they **hear**
- 30% of what they **see**
- 50% of what they **hear and see**
- 70% of what they **say and write**
- 90% of what they **do**

Therefore, the learning outcomes that educators and instructors set for their courses need to coincide with what students experience — if we want our students to remember what they are being taught then we must allow them to actively do things (e.g., active learning activities) and provide multiple opportunities for this type of learning.

Felder and Brent (1996) outlined that student-centred instruction is a teaching approach that incorporates active learning into lectures, holds students responsible for their learning, and utilizes self-paced and/or cooperative (i.e., team or group based) learning. Although the promised benefits of student-centred instruction are real it is crucial to recognize that they are neither immediate or automatic. It takes time for these approaches to "sink in" and may require repetitions. For example, having students write down a Minute Paper at the end of the class where they identify the most significant things they have learned from a specific lecture, discussion, or argument may not be helpful to them if it is done only once in the term. Rather, it may be necessary to have students write down a Minute Paper at the end of each section or at the end of each week so that they continually check-in with themselves using this type of low-stakes formative assessment.

Active learning activities are important and necessary even though they may seem daunting or difficult for different learning environments (e.g., large enrollment courses, small classes or online courses). Active learning creates an intellectually stimulating and, at times, challenging learning environment regardless of the type of environment. "The more that students became active partners in the learning process, the more they took ownership of the course and of their learning" (Ebert-May et al., 1997, p. 606). For instance, small-group activities in large classes allow students to interact with their peers, hear other perspectives, decrease one's anonymity in a large group setting, and encourages engagement with course material (Yazedjian and Kolkhorst, 2007). Using active learning strategies in various learning environments is beneficial but it must be accompanied by:

1. An explanation of the purpose of the activity, either beforehand or during the debrief once the activity has ended;

2. Explicit directions of the tasks or steps that the students are required to perform; and

3. A schedule outlining the time they have for each task or step.

A discussion of the results or conclusions of the activity may also occur between the entire class depending on the type of activity, class environment, or time.

Past research on active learning activities in the fields of mathematics and statistics must also be considered. Implementing cooperative learning using small groups in introductory statistics courses increases student's success rate of passing the course, increases their course marks and generally increases their satisfaction in the course (Keeler and Steinhorst, 1995; Garfield, 1993). Rosenthal (1995) discussed both formal and informal written assignment activities (e.g., Minute Paper) in the mathematics classroom to encourage students to think more deeply about course content. Kerrigan (2018) provided three quick fixes for making undergraduate mathematics courses more active and engaging, such as running class polling at the beginning of a class to elicit prior knowledge as part of a pre-assessment review of concepts. Integrating active learning activities into theoretical undergraduate statistics courses is also possible (Prins, 2009), whereby these activities help to create a supportive community and reinforce learning of course concepts. Carlson and Winquist (2011) evaluated the effectiveness of a semester-long active learning statistics curriculum involving workbooks, resulting in positive changes in students' attitudes towards statistics which are positively associated with performance on the comprehensive final exam.

Michael (2006) outlined that there is evidence that student-centred active learning strategies really work, stating that:

> "We should all begin to reform our teaching, employing those particular approaches to fostering active learning that match the needs of our students, our particular courses, and our own teaching styles and personalities. There are plenty of options from which we can choose, so there is no reason not to start. This will mean that we too become learners in the classroom." (p. 165)

Therefore educators must become life-long learners themselves who, hopefully, take risks in the classroom by experimenting with active learning strategies to determine which ones work well in their courses and for their students. "At this point it is unethical to teach any other way" (Waldrop, 2015, p. 273).

Chapter 3

Characterizing Two Phases of Wildland Fire Lifetimes with Multi-State Models

As noted earlier, past research on wildland fire lifetimes generally only considered one subcomponent or *phase* of a fire's evolution. The objectives of our research on fire lifetime phases is to determine which factors influence the length of time spent in each phase and how the phases are interrelated. We begin by diving deeper into the context of our wildland fire dataset since it is an important part of advancing our understanding of the methods developed, utilized, and interpreted in this chapter and Chapter 4.

3.1 Study Area

We utilize a wildland fire dataset from the Sioux Lookout District in Ontario, consisting of 2,239 fully suppressed fires that occurred between 1989 and 2019. These data were provided by the Ontario Ministry of Northern Development, Mines, Natural Resources and Forestry (hereafter referred to as the Ministry). Figure 3.1 plots these fires which are located in the Ministry's Northwest Fire Region.

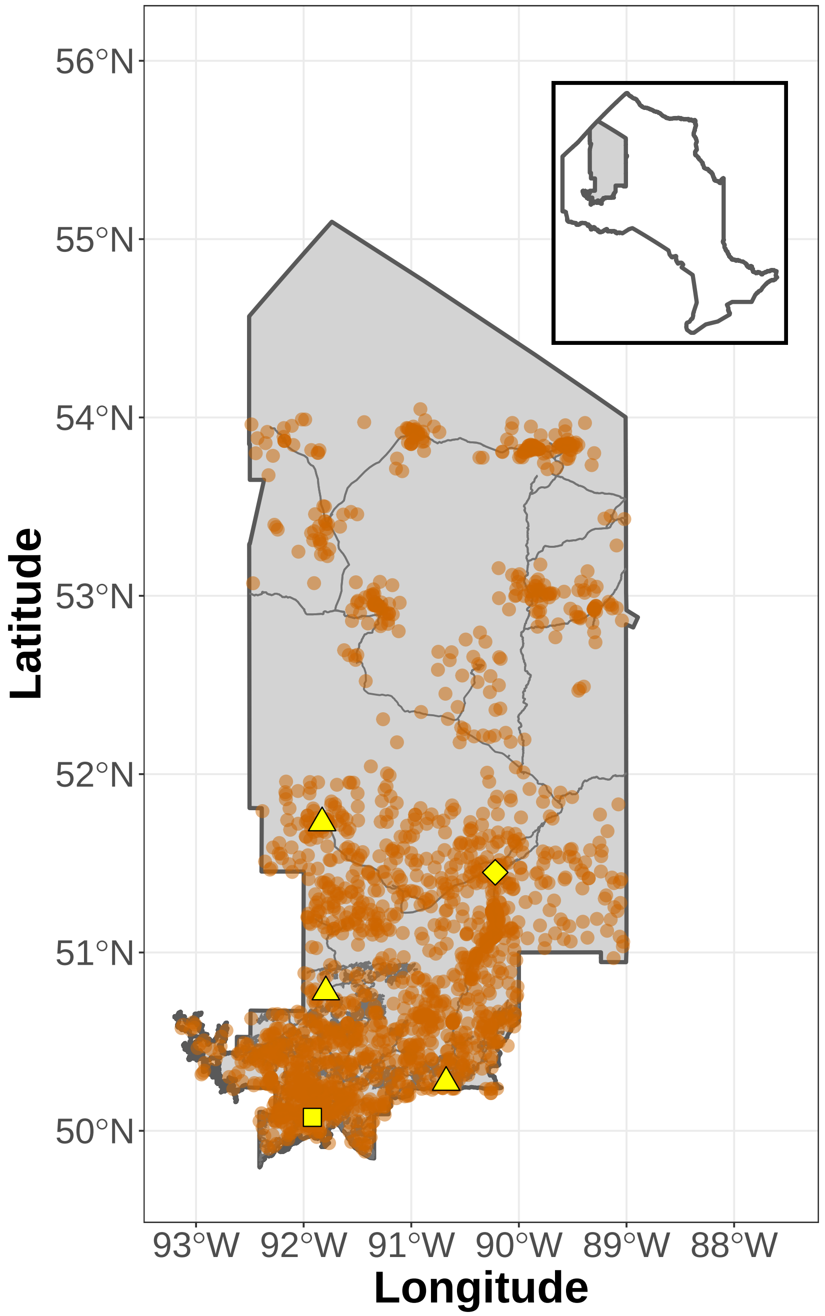


Figure 3.1: Map of Sioux Lookout District in Ontario (see inset) showing the fire locations (orange circles). Roads (dark grey lines) and important fire management locations such as the Headquarters (yellow square), Attack Base (yellow diamond), and Forward Attack Bases (yellow triangles) are also shown.

Fire management agencies like the Ministry must decide how they manage wildland fires across their jurisdictions. In the past the Wildland Fire Management Strategy (OMNR, 2004) used by the Ministry for responding to such fires included: full suppression, partial suppression, and monitoring of a fire. Full suppression required that either direct or indirect action be taken on the entire perimeter to acquire control, whereas partial suppression required action to be taken on key areas of the fire perimeter. Monitored fires were observed and continually assessed to determine if further response was required (i.e., to minimize social disruptions and/or economic impacts).

In 2015, Ontario implemented its new Wildland Fire Management Strategy (OMNRF, 2014b) which changed to a system of *appropriate response*, whereby "each fire is assessed and receives an appropriate response based on the circumstances and condition of the fire". Fire management practitioners in Ontario must assess the potential impact of every fire that is reported and decide whether it should be monitored or managed with some form of modified suppression action. This modified approach to fire response recognizes that fire is a natural process that provides ecological benefits to forest ecosystems while maintaining safety by minimizing damage and distruptions caused by fire.

Importantly, this fire response alteration occurred during the time period of our data and so we must be mindful of it when drawing conclusions. We have restricted our dataset to fully suppressed fires, even for the fires from 2015-2019, as our interest is in fires that progress through the various phases, as discussed later in this chapter. Approximately 60% of these fires were started by lightning. The remaining 40% were ignited by humans. There are a variety of different ways people can ignite fires, such as through recreational activity, by residential activity (e.g., clearing land), railways and industrial forestry activities, as key examples. Figure 3.2 shows the number of fires in each year stratified by their cause of ignition (human vs. lightning). The number of human- and lightning-caused fires in Sioux Lookout appears to fluctuate every few years with the latter exhibiting greater changes over the years. Model fitting and analyses in this chapter and Chapter 4 are separated based on the type of fire cause — human and lightning — as was done in other analyses using survival analysis methods for fires in Ontario (Morin et al., 2015, 2019).

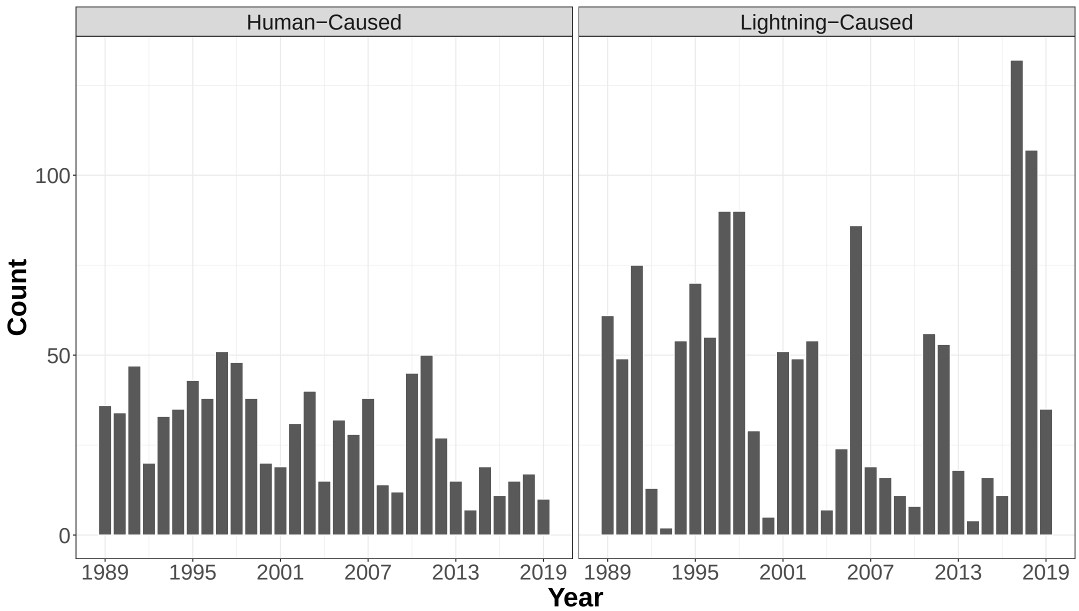


Figure 3.2: Bar plot of the fires, showing the number of fires in each year by cause.

3.2 Data

The lifetime of a fire, shown visually in Figure 3.3, is split up into several key events that are of interest. Our Sioux Lookout data contains date (yyyy-mm-dd) and time (hr-min) stamps for the following events (OMNRF, 2014a):

- The *ignition* or *start* of the fire which is either known or estimated (cf. Figure 3.4),
- The *discovery* of the fire,
- When it is *reported* (first and second instances),
- The *getaway* time: when resources are dispatched to a fire,
- The time that the *initial attack* on the fire began: by suppression efforts (e.g., air tankers and ground forces, discussed later),

- The time when it was declared to be in a condition of *being held*: with currently committed resources, sufficient suppression action has been taken so the fire is unlikely to spread beyond existing control boundaries under forecasted conditions,
- The time when it was declared to be in a condition of *under control*: the fire has received enough suppression action to ensure no further spread of the fire, and
- The *out* time: the time that the fire was declared to be extinguished.

Two *phases* of interest are highlighted in Figure 3.3. The **detection phase**, encompassing the start of a fire to when it's reported, and the **action phase**, a combination of how long a fire is in the system, the travel time to that fire, and the time it takes to bring a fire under control.

Figure 3.3: Visual representation of fire lifetime events, along with the detection phase (orange) and the action phase (blue).

3.3 Methods

Multi-state models are frequently used in epidemiology and biostatistics to model the life history of an individual. One common example is the illness-death model, shown in Figure 3.5, where individuals in a study are in a healthy state at the start and may transition to the two other states 'sick' and 'dead'. If a person gets sick with a specific disease of interest then they move to the 'sick' state. People can either die when they are healthy or when they are sick with the disease. Individuals in this type of study may be censored if they do not get sick or die during the study period since it is often infeasible to follow an individual for their entire lifetime.

Now suppose that individuals only die if they become ill from a disease and once they do they do not return to the healthy state but progress to the death state. Then the

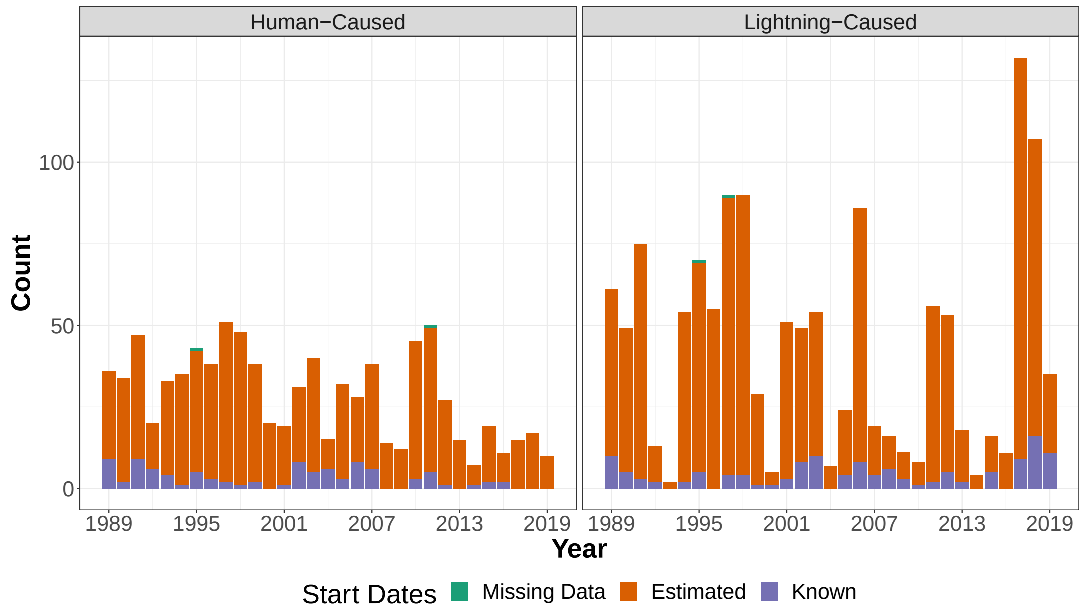


Figure 3.4: Stacked bar plot of the fires by cause for each year, showing the number of start dates that were known ($\approx 10\%$), estimated ($\approx 90\%$), and missing ($< 0.2\%$) across all fires.

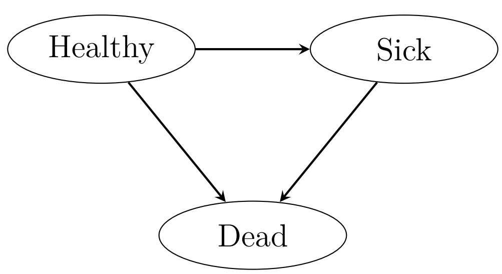


Figure 3.5: Diagram of the simple illness-death model.

illness-death model is called a 3-progressive process since it progresses solely through the three states as illustrated in Figure 3.6. Here, there may still be censoring if an individual is never observed to get sick or die by the end of the study period.

Figure 3.6: Diagram of a 3-progressive illness-death model.

Recall that Morin et al. (2015) modelled the control time of fires in Ontario using a single fire lifetime phase. They defined the *control time* of a fire to be the time interval from the start of initial attack action to the time that the fire is declared as being under control. This survival model can be considered as a 2-phase progressive model from initial attack to under control. In this thesis, the lifetime process that we utilize for fully suppressed human- and lightning-caused fires is a 3-progressive process with two lifetime phases of interest (detection and action). Figure 3.7 illustrates this process. One important difference between our process and the illness-death one is that no censoring takes place for our observations since the entire lifetime of our fires are observed.

Figure 3.7: Diagram of our 3-progressive multi-state model for fire lifetimes.

3.3.1 Exploratory Analysis

Ninety fully suppressed fires (51% human-caused; 49% lightning-caused) were removed from the dataset since the detection and action phases: (1) could not be calculated due to missing values; (2) had negative phase lengths due to the preceding event likely being incorrectly recorded after the subsequent event; or (3) had zero-length phases due to the events being recorded at the same time. We investigated the fires with zero-length phases and found no clear pattern or trend as to why they occurred. A zero-length detection phase fire could occur because a fire is seen and reported simultaneously during a loaded patrol, or firefighters working on a rail fire see a train spark another fire along the track, or firefighters working on a fire in the community encounter people starting another fire while they are there. Similarly, a zero-length action phase fire could occur because municipal or industry personnel reported a fire that they themselves had actioned making the time instances the same. After also removing fires with missing variables, where the details about such variables are provided in Section 3.3.2, we utilize 786 human-caused fires and 1,270 lightning-caused fires from the Sioux Lookout District, approximately 92% of the original data, for our work.

Figure 3.8 shows the distributions of the phase lengths for the two fire causes while Table 3.1 provides the minimum and maximum phase lengths. For Figure 3.8, the scale of the x-axis is the fire phase lengths in days (cut-off at 4 days) and the scale of the y-axis is the density of those fires; the inset provides a histogram of the fire lifetimes where the scale of the x-axis is the logarithmic transformation (base 10) of the fire phase lengths and the scale of the y-axis is the counts of those fires. Lightning-caused fires have longer detection and action phases than human-caused ones, unsurprisingly, since lightning-caused fires can often go undetected for days or can be smouldering underground, and because they can occur in remote areas that may be more difficult for suppression crews and equipment to access. Approximately 90% of human-caused fires in Sioux Lookout are detected and undergo action within 24 hours of ignition.

3.3.2 Variables of Interest

The data available for each fire observation and considered for model fitting, categorized by homogeneous sub-groups, are described in Table 3.2. The fuel moisture variables from the Canadian Fire Weather Index (FWI) System are functionally related in a hierarchical structure as shown in Figure 3.9. These variables account for the effects of fuel moisture and weather conditions on fire behaviour. The calculation of these variables is

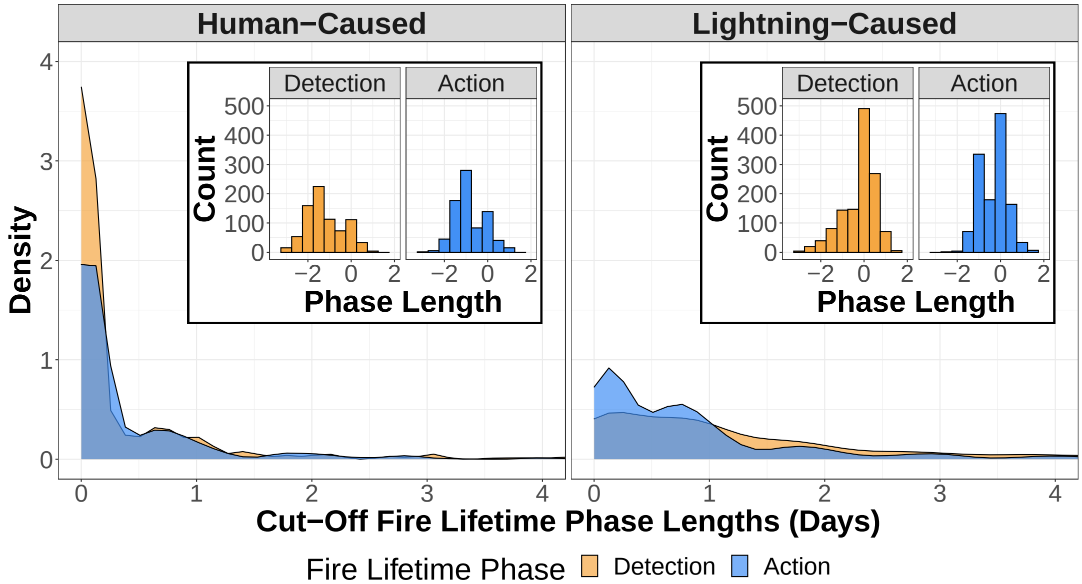


Figure 3.8: Distributions of the fire lifetime phase lengths for human- and lightning-caused fires. The inset histograms are of the transformed fire phase lengths in $\log_{10}$ scale.

Table 3.1: Maximum and minimum fire lifetime phase lengths.

Fire Cause & Fire Phase	Minimum	Maximum
Human		
Detection	1 minute	13.9 days
Action	1 minute	16.0 days
Lightning		
Detection	1 minute	32.7 days
Action	6 minutes	65.1 days

based on daily observations of temperature, relative humidity, wind speed, and 24-hour precipitation. The Sioux Lookout District data consists of single numeric ratings of each of these variables for every fire; these ratings are the observed values on the day a fire was ignited. We take great care when using these variables in our models because of potential multicollinearity. The Ministry uses the variable ratings as a guideline of the relative potential for wildland fire. See Van Wagner (1987) and Wotton (2009) for more details about the FWI system and its components.

The simple fuel type variables accounts for the type of fuel (i.e., vegetation) feeding the fire at ignition. Several temporal, distance, and other variables are considered as well. Seasonality is measured by the meteorological seasons, in this case Spring includes March, April, and May, Summer includes June, July, and August, and Fall includes September, October, and November. The early ignition and early report indicator variables were created to determine if the time of day — early (i.e., before noon) or late (i.e., after noon) — that a fire is ignited or reported affects the detection or action phase lengths, respectively. Morin et al. (2015) used a similar variable to account for the time of day when the initial attack of a fire started.

We define the *successful initial attack* of a fire as being declared "being held" by end of day the day after a fire was reported, and less than or equal to 100 hectares (ha), or less than or equal to 4 ha with no limit on the time to being held. Therefore this definition encompasses very small fires that are easy to suppress and larger fires where the suppression has been within a short period of time (2 days). This definition was developed in collaboration with a Forest Fire Science Specialist from the Ministry.

Most of the continuous variables have been standardized by subtracting the mean and dividing by the standard deviation across all fires within each category of human- and lightning-caused fires to allow for easier interpretations. The standardized variables include: all fuel moisture variables, all distance variables, fire load, and initial attack size. Number of air tankers, number of ground forces, and detection phase lengths were not standardized.

Table 3.2: Overview of the data and variables considered when modelling. A superscript of 1 or 2 implies that these variables are considered when modelling the detection or action phases, respectively.

Variable	Description
Fuel Moisture[1,2]	
Fine Fuel Moisture Code (FFMC)	A numeric rating of the moisture content of litter and other cured fine fuels
Duff Moisture Code (DMC)	A numeric rating of the average moisture content of loosely compacted organic (duff) layers of moderate depth
Drought Code (DC)	A numeric rating of the average moisture content of deep, compact organic layers
Initial Spread Index (ISI)	A numeric rating of the expected rate of fire spread, which combines the effects of wind speed and FFMC
Build-Up Index (BUI)	A numeric rating of the total amount of fuel available for combustion, based on the DMC and DC
Fire Weather Index (FWI)	A numeric rating of fire intensity based on the ISI and BUI, providing a general index of fire danger in forested areas of Canada
Fuel Type[1,2]	
Grass Fuel	An indicator of whether grass fuel types (1) were burning and contributing most to the forward spread of the fire at the time of initial response
Coniferous Fuel	An indicator of whether coniferous fuel types (1) were burning and contributing most to the forward spread of the fire at the time of initial response
Mixedwood Fuel	An indicator of whether mixedwood fuel types (1) were burning and contributing most to the forward spread of the fire at the time of initial response
Other Fuel	An indicator of whether other fuel types (1) were burning and contributing most to the forward spread of the fire at the time of initial response, used as the baseline
Temporal	
Spring[1,2]	An indicator of whether the fire was ignited in Spring (1)
Summer[1,2]	An indicator of whether the fire was ignited in Summer (1), used as the baseline
Fall[1,2]	An indicator of whether the fire was ignited in Fall (1)
Early Ignition[1]	An indicator of the time of day when a fire is ignited, either early (1) or late (0)
Weekday Ignition[1]	An indicator of whether the fire was ignited on a weekday (1) or weekend (0)

Variable	Description
Temporal	
Early Report[2]	An indicator of the time of day when a fire is reported, either early (1) or late (0)
Weekday Report[2]	An indicator of whether the fire was reported on a weekday (1) or weekend (0)
Same Day Detection[2]	An indicator of whether the ignition and report of a fire occur on the same day (1)
Same Day Dispatch[2]	An indicator of whether the report of and getaway to a fire occur on the same day (1)
Distance[1,2]	
Distance to FMH	The distance (in km) from a fire to the closest Fire Management Headquarters (FMH) in Sioux Lookout
Distance to AB	The distance (in km) from a fire to the closest Attack Bases (AB) in Sioux Lookout
Distance to FAB	The distance (in km) from a fire to the closest Forward Attack Bases (FAB) in Sioux Lookout
Distance to Road	The distance (in km) from a fire to the nearest road in Sioux Lookout
Other[2]	
Fire Load	The number of fires burning on the landscape in Sioux Lookout at the same time as a given fire
Initial Attack (IA) Size	The size of a fire (in hectares) at the time of initial attack
Ground Forces	The size of the ground forces (e.g., fire fighters) at initial attack
Air Tankers	The initial number of air tankers used on the fire
Successful Initial Attack (IA)	An indicator of whether the initial attack of a fire was deemed successful (1) or unsuccessful (0)
Detection	The duration (in days) of the detection phase

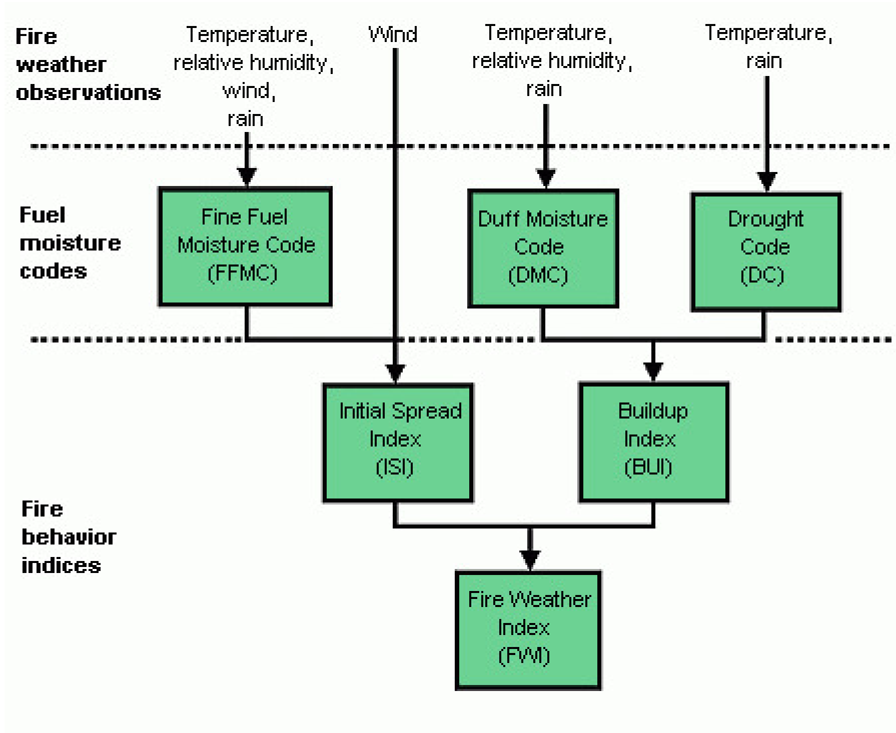


Figure 3.9: Structure of the FWI System. Copied from the Natural Resources Canada website, https://cwfis.cfs.nrcan.gc.ca/background/summary/fwi.

3.3.3 Modelling Framework

The multi-state models for human- and lightning-caused fires that follow the process shown in Figure 3.7 have state space $\{1, 2, 3\}$ where states 1, 2, and 3 represent the ignition, report, and under control states, respectively. We model the lifetime process where predictors act multiplicatively on the intensities of transitions between states resulting in a framework that is similar to the Cox PH models. For fire i, $i = 1, \ldots, n$, the transition intensity functions take the following form

$$\begin{aligned}\lambda_{i,kl}(t|\mathcal{H}(t^-)) &= \lambda_{i,kl}(t|\boldsymbol{x}_i, \mathcal{H}(t^-)) \\ &= \lambda_{kl0}(t|\mathcal{H}(t^-)) \exp(\boldsymbol{x}_i' \boldsymbol{\beta}_{kl}),\end{aligned} \tag{3.1}$$

where $k, l \in \{1, 2, 3\}$ index the transition from state k to state l, $l > k$, $\boldsymbol{x}_i$ is a column vector of the predictors for fire i, $\boldsymbol{\beta}_{kl}$ is a column vector of the corresponding fixed effects coefficients, and $\lambda_{kl0}(t|\mathcal{H}(t^-))$ is the unspecified baseline transition intensity function. The value t refers to either the detection or action phase length depending on the transition that is taking place.

There are two ways to treat time in multi-state modelling: the clock-forward and clock-reset approaches. For *clock-forward* models, time is measured from the initial phase, whereas for *clock-reset* models the clock is set back to zero every time a new phase is entered, therefore time is only measured from the start of the new phase (Jackson, 2016; Williams et al., 2017). Clock-forward models are Markov models since the movement from the current phase does not depend on history, but for clock-reset models the timescale does depend on history so these models are semi-Markov instead of Markov. When t is set as the clock-reset time then the history, $\mathcal{H}(t^-)$, is present in the model for (3.1) whereas it is not when t is set as the clock-forward time.

The coefficients $\boldsymbol{\beta}_{kl} = (\beta_{kl1}, \beta_{kl2}, \ldots, \beta_{klm})'$ measure the impact or the effect size of the m predictors for the k to l transition. If the p^{th} predictor is a dichotomous variable that has only two levels or categories, such as the early ignition predictor, then the ratio of the transition intensity functions for those two levels is given by the hazard ratio (HR) $\exp(\beta_{klp})$, assuming that everything else is constant. If the p^{th} predictor is continuous then the ratio of the transition intensity functions is given by the HR $\exp(\beta_{klp})$ for a unit difference of that predictor (e.g., a unit increase in the predictor), assuming that everything else is constant. For instance, suppose β_{klp} is positive then the HR is greater than one. This implies that as the value of the p^{th} predictor increases, the hazard function also increases resulting in a faster decrease of the survival function. The opposite relationship holds when β_{klp} is negative and the HR is less than one. This implies that as the value of the p^{th} predictor increases, the hazard function decreases resulting in a slower decrease of the survival function. There is no effect when β_{klp} equals zero. In the fire lifetime context, if the survival function decreases faster (slower) then the length of survival for fires in a specific phase will decrease faster (slower) and the phase will not last as long (will last longer).

3.3.4 Considerations Prior to Modelling

Before fitting any multi-state models we fit two separate Cox proportional hazards (PH) models to the detection and action phases using the `survival` package (Therneau, 2020). This exploratory modelling was conducted to help identify which predictors may be important for each phase. All of the predictors listed in Table 3.2 were included in the models. We used stepwise model selection by AIC from the `MASS` package (Venables and Ripley, 2002) which compares the AIC improvements when using both forward and backward selection. These standard statistical modelling procedures were used in conjunction with our domain knowledge of fire science and management and exploratory data analyses to determine which predictors to include in the subsequent multi-state models. A summary of these predictors are shown in Table 3.3.

Fitting multi-state models to completely-observed data where one knows the complete history of the process may be achieved through the use of the `survival` and `mstate` (de Wreede et al., 2011) packages in R and converting the data into so-called long format where there is one line for each transition. The Sioux Lookout data provided by the Ministry was in wide format with one subject (i.e., fire) per line. Converting the data from wide format to long format is outlined in Appendix A. See Putter et al. (2007) for more details.

We begin with defining the possible transitions of the 3-progressive process by specifying a transition matrix where direct transitions are possible (those with NA, meaning not applicable, are impossible) and assigning numbers to these transitions for future

Table 3.3: Summary of predictors (fixed effects) included in each model, represented by a checkmark, in the suite of human- and lightning-caused models for Ontario's SLK District.

	Human-Caused		Lightning-Caused	
Predictor	Detection	Action	Detection	Action
FFMC			✓	✓
DMC			✓	
FWI	✓			
Mixedwood Fuel	✓	✓	✓	✓
Grass Fuel	✓	✓	✓	✓
Other Fuel	✓	✓	✓	✓
Spring	✓	✓		
Fall	✓	✓		
Early Ignition	✓		✓	
Weekday Ignition	✓			
Early Report				✓
Same Day Dispatch		✓		✓
FMH Distance		✓		✓
AB Distance	✓	✓		
FAB Distance	✓			
Road Distance	✓			✓
Fire Load				✓
IA Size				✓
Ground Forces		✓		✓
Air Tankers		✓		✓
Successful IA		✓		✓
Detection		✓		✓

reference. The transition matrix for the 3-progressive process is

$$
\begin{array}{cc}
 & \begin{array}{ccc} \mathit{IGN} & \mathit{REP} & \mathit{UCO} \end{array} \\
\begin{array}{c} \mathit{IGN} \\ \mathit{REP} \\ \mathit{UCO} \end{array} &
\begin{pmatrix} NA & 1 & NA \\ NA & NA & 2 \\ NA & NA & NA \end{pmatrix}
\end{array}
$$

where IGN, REP, and UCO are abbreviations of the three states ignition, report, and under control, respectively. We see that transition 1 from ignition to report represents the detection phase of the fire lifetime and transition 2 the action phase.

For the human-caused fires, there are 116 detection times that are tied meaning that two or more human-caused fires have the same detection time whereas there are 127 uniquely tied action times. For instance the detection time of 5 minutes occurs for 28 human-caused fires. Overall, approximately 73% of the detection times and 64% of the action times are tied for human-caused fires. These percentages are lower for lightning-caused fires (35% detection and 50% action). The methods for handling ties mentioned in Section 2.2 are nearly equivalent if the data contains very few ties which is not the case for our Sioux Lookout data. We utilize Efron's approximation to deal with the tied fire lifetime phase times as suggested by Hertz-Picciotto and Rockhill (1997).

3.4 Analysis and Results

Both the clock-reset (CR) and clock-forward (CF) time multi-state models were fit to the human- and lightning-caused fire data. Figures 3.10 and 3.11 provide plots of the Kaplan-Meier estimates of the survival functions for multi-state *null* models fit without any predictors to examine the survival curves of the two phases. It is important to note that the proportional hazards assumption may be violated since the survival probabilities for the two phases cross one another for the CF human-caused multi-state null model and the CR lightning-caused multi-state null model.

The survival curves of the two phases have a similar pattern of quickly decreasing to zero over time. We can see that the Ministry is generally quick at detecting and actioning human-caused fires since their survival curves are relatively close together. However, the detection and action survival curves appear to be farther apart from one another for the lightning-caused fires, with the detection curve having a higher probability of survival than the action curve within the first ten days of a fire, suggesting that the Ministry is slower at detecting lightning-caused fires. This noticeable smooth exponential decay of the detection curve for lightning-caused fires was previously mentioned in Wotton and Martell (2005).

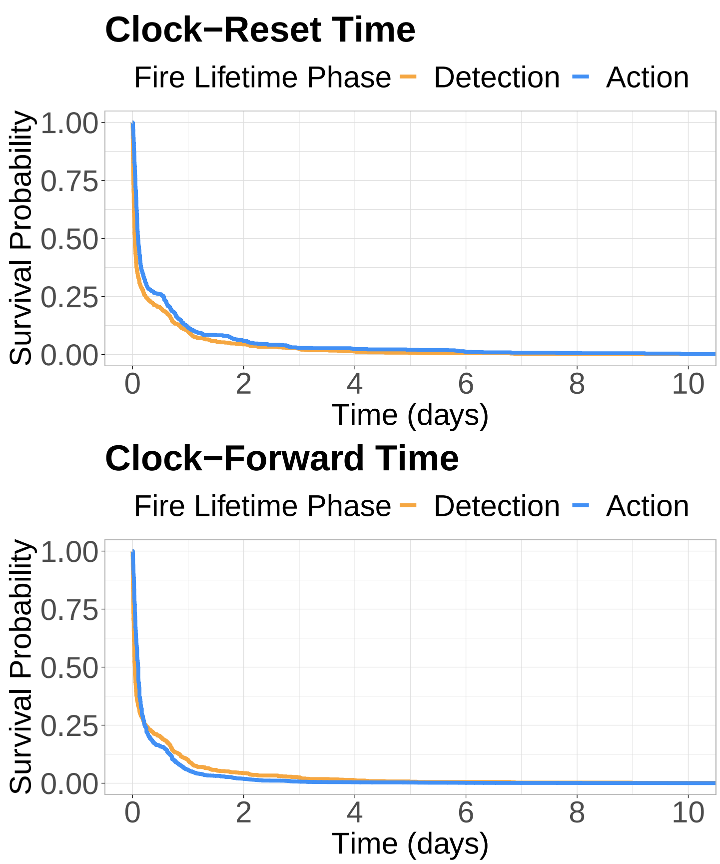


Figure 3.10: Plots of the survival functions for the human-caused clock-reset and clock-forward multi-state null models.

Tables 3.4 and 3.5 provide the HR estimates and standard errors of the fixed effect

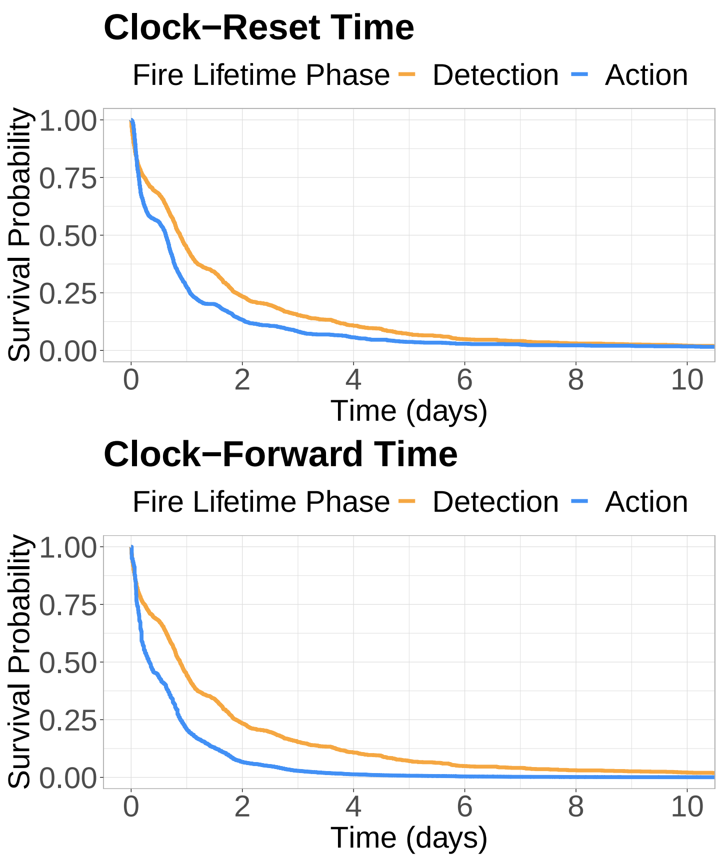


Figure 3.11: Plots of the survival functions for the lightning-caused clock-reset and clock-forward multi-state null models.

coefficients associated with the predictor variables for the four models. To better understand the interpretations of the HR estimates, consider the baseline case where we have a single predictor. A HR greater than one implies that the length of survival is decreasing (i.e., a shorter phase length) whereas a HR less than one implies that the length of survival is increasing (i.e., a longer phase length). In the multiple predictor situation, these interpretations hold for a given predictor when all other predictors are held constant.

Evidence for a HR estimate is assessed using the Wald statistic, corresponding to the ratio of the fixed effect coefficient to its standard error, which evaluates whether the coefficient of a given variable is statistically significantly different from zero. HR estimates with one, two, or three stars indicate weak (i.e., p-value < 0.1), moderate (i.e., p-value < 0.05), or strong (i.e., p-value < 0.01) evidence, respectively. Many of the HR estimates have a minimum of two stars indicating that there is moderate to strong evidence that the corresponding predictors have an effect on their respective phases. The interpretations of the predictors with moderate and strong evidence from the **human-caused** multi-state models in Table 3.4 follow:

- **FWI:** Fires burning under more intense conditions are associated with a shorter detection phase. This may be due to the Ministry and communities being on a higher alert for such fire danger. Also, a fire burning under more intense conditions may be emitting more smoke and therefore would be easier to detect.

- **Fuel:** Fires burning from mixedwood fuel have longer detection phases than those burning from coniferous fuel, while fires burning from grass fuel have shorter detection and action phases than those burning from coniferous fuel. Grass fires are often easier to detect because they occur near people and there are no tree canopies blocking them. As well, grass fires may be easier to contain because they are easier to access and it is often easier to work in this fuel type.

- **Season:** Fires ignited in spring have shorter detection phases than those in summer. "Green-up" effects occur in spring where the new cycle of plant growth begins, most likely allowing for an easier detection of fires. Fires ignited in fall have longer detection and action phases than those in summer. During the fall many detection aircraft go off contract and there is less sunlight during the day, potentially contributing to a longer detection. Also, fewer fire crews work in the fall season and fires may be deemed as less urgent due to mild fire conditions which could contribute to longer action phases.

- **Time of Ignition:** Fires ignited earlier in the day have a longer detection phase than those ignited after noon while fires ignited during a weekday have a shorter detection phase than those ignited on the weekend. Detection patrols start flying by the late morning or early afternoon; therefore fires ignited prior to noon occur when detection patrols are not running and thus take longer to be detected. Human-caused fires include fires started by industrial forestry activities which typically operate during weekdays; therefore fires ignited during a weekday may be detected faster because they are ignited and reported by industry.

- **Same Day Dispatch:** Fires that have same day dispatch tend to have a much shorter action phase than those that do not. This may be due to the fact that the fire does not have much time to spread and become unmanageable from the time it is reported. Or if the reported fire is already large, then it does not have time to become worse.

- **Distance:** Human-caused fires located farther away from the AB or roads are associated with a longer detection phase. Note that the public finds and reports many fires. If these fires are not near areas of higher population density then they may take longer to be detected. However, human-caused fires located farther away from the FAB are associated with a shorter detection phase. This may be due to the locations of the FAB in Sioux Lookout and that they are only staffed when necessary. Similarly, human-caused fires located farther away the FMH or AB are associated with a longer action phase. In Sioux Lookout, the FMH, AB, and FAB are all clustered mainly in the south, allowing for vast remote regions without nearby fire management resources.

- **Suppression Efforts:** Fires that require more suppression efforts are associated with a longer action phase. This makes sense because a fire that requires more initial resources to get it under control would likely take longer to suppress, resulting in a longer action phase.
- **Successful IA:** Fires that have a successful initial attack have a much shorter action phase than those that do not.
- **Detection:** Notably the detection phase length shows no evidence that it has an effect on the action phase length for the CR model for human-caused fires.

The interpretations of the predictors with moderate and strong evidence from the **lightning-caused** multi-state models in Table 3.5 follow:

- **FFMC & DMC:** Fires burning under drier conditions of the smallest forest fuels and medium-sized fuels are associated with a shorter detection phase. Whereas fires burning under drier conditions of only the smallest forest fuels are associated with a longer action phase. This makes sense because as conditions become more dry, more fuels can burn and the faster a fire can spread, resulting in more smoke which aids in detection and may result in a fire that is more challenging to suppress.
- **Fuel:** Fires burning from mixedwood fuel have longer detection phases than those burning from coniferous fuel, whereas fires burning from other fuel types have shorter action phases than those burning from coniferous fuel. In discussions with our fire science collaborators, we have no clear reasons for this occurrence, although

we suspect it may be related to seasonality and locations of these fuels. Further investigation is required.

- **Time of Ignition & Report:** Fires ignited earlier in the day have a longer detection phase than those ignited after noon. Fires reported earlier in the day have a longer action phase than those reported after noon. If a lightning-caused fire is reported earlier in the day then it may be more intense in terms of size and severity, aiding in its detection and resulting in a longer action phase.

- **Same Day Dispatch:** Fires that have same day dispatch tend to have a much shorter action phase than those that do not.

- **Distance:** Lightning-caused fires located farther away from the FMH or roads are associated with a longer action phase. Fires that are located in remote areas can take longer for crews to travel to, allowing the fire to grow more prior to starting initial attack. As well, fires burning in remote areas can be harder to access and therefore harder to fight.

- **Fire Load & IA Size:** Fires burning when there are more fires already burning on the landscape are associated with a longer action phase. Also, fires with a larger initial attack size are associated with a longer action phase. Both of these scenarios make sense because if there are more fires burning on the landscape then suppression efforts may be stretched thin, resulting in a longer action phase. Similarly, the larger a fire is at initial attack then the longer it will take to bring it under control because it will require more suppression efforts.

- **Suppression Efforts:** Fires that require more suppression efforts are associated with a longer action phase.

- **Successful IA:** Fires that have a successful initial attack have a much shorter action phase than those that do not.

- **Detection:** Fires with longer detection phases are associated with longer action phases for the CR model.

Table 3.4: Summary of hazard ratio estimates (standard errors of fixed effects coefficients) for the human-caused multi-state models.

	Detection Phase		Action Phase	
Predictor	Clock-Reset Model	Clock-Forward Model	Clock-Reset Model	Clock-Forward Model
FWI	1.09** (0.04)	1.09** (0.04)		
Mixedwood Fuel	0.78** (0.11)	0.78** (0.11)	0.94 (0.11)	0.93 (0.11)
Grass Fuel	1.48*** (0.10)	1.48*** (0.10)	1.78*** (0.10)	1.65*** (0.11)
Other Fuel	1.08 (0.11)	1.08 (0.11)	1.13 (0.11)	1.10 (0.11)
Spring	1.43*** (0.09)	1.43*** (0.09)	0.96 (0.09)	0.91 (0.09)
Fall	0.49*** (0.14)	0.49*** (0.14)	0.55*** (0.14)	0.50*** (0.15)
Early Ignition	0.66*** (0.09)	0.66*** (0.09)		
Weekday Ignition	1.19** (0.08)	1.19** (0.08)		
Same Day Dispatch			2.63*** (0.16)	2.90*** (0.17)
FMH Distance			0.87*** (0.04)	0.84*** (0.04)
AB Distance	0.87*** (0.04)	0.87*** (0.04)	0.91** (0.04)	0.92** (0.04)
FAB Distance	1.14*** (0.04)	1.14*** (0.04)		
Road Distance	0.71*** (0.08)	0.71*** (0.08)		
Ground Forces			0.98* (0.01)	0.98** (0.01)
Air Tankers			0.55*** (0.06)	0.54*** (0.06)
Successful IA			6.26*** (0.29)	5.04*** (0.27)
Detection			0.93 (0.05)	

Note: *p<0.1; **p<0.05; ***p<0.01

Table 3.5: Summary of hazard ratio estimates (standard errors of fixed effects coefficients) for the lightning-caused multi-state models.

	Detection Phase		Action Phase	
Predictor	Clock-Reset Model	Clock-Forward Model	Clock-Reset Model	Clock-Forward Model
FFMC	1.13*** (0.03)	1.13*** (0.03)	0.94** (0.03)	0.95* (0.03)
DMC	1.08** (0.03)	1.08** (0.03)		
Mixedwood Fuel	0.82** (0.08)	0.82** (0.08)	0.86* (0.09)	0.88 (0.9)
Grass Fuel	1.01 (0.36)	1.01 (0.36)	1.17 (0.36)	1.21 (0.36)
Other Fuel	1.03 (0.09)	1.03 (0.09)	1.52*** (0.10)	1.51*** (0.10)
Early Ignition	0.79*** (0.06)	0.79*** (0.06)		
Early Report			0.82** (0.09)	0.90 (0.09)
Same Day Dispatch			2.30*** (0.09)	2.18*** (0.09)
FMH Distance			0.80*** (0.04)	0.82*** (0.03)
Road Distance			0.94** (0.03)	0.93*** (0.03)
Fire Load			0.94** (0.03)	0.95* (0.03)
IA Size			0.94** (0.03)	0.95* (0.03)
Ground Forces			0.94*** (0.01)	0.95*** (0.01)
Air Tankers			0.67*** (0.04)	0.69*** (0.04)
Successful IA			4.04*** (0.14)	3.36*** (0.13)
Detection			0.96*** (0.01)	

Note: *p<0.1; **p<0.05; ***p<0.01

Table 3.6 provides measures of the model fit for all of the multi-state models. The Wald and likelihood ratio significance tests compare the fitted model to a null model (i.e., intercept-only model) which predicts the mean survival time for all fires. The logrank test compares the survival distributions of the fitted model to the null model. All of the tests show that the four fitted models are preferred over the intercept-only model.

Table 3.6: Measures of model fit for multi-state models. The human-caused models have 23 and 22 degrees of freedom for the clock-reset and clock-forward models, respectively. The lightning-caused models have 20 and 19 degrees of freedom for the clock-reset and clock-forward models, respectively.

	Wald Test	Likelihood Ratio Test	Score (Logrank) Test
Human-Caused			
Clock-Reset Model	538.10***	670.12***	577.75***
Clock-Forward Model	566.36***	742.17***	626.77***
Lightning-Caused			
Clock-Reset Model	673.70***	899.88***	693.81***
Clock-Forward Model	710.27***	1,033.83***	864.43***

Note: *p<0.1; **p<0.05; ***p<0.01

Figures 3.12 and 3.13 show the survival functions for the fitted CR and CF multi-state models for human- and lightning-caused fires, respectively. For the human-caused fire models the survival probability for the detection phase quickly decreases to zero over four days while the action phase steadily decreases over ten days. Therefore, the probability that a human-caused fire is still not completely under control approximately zero to ten days after it was detected is higher than the probability that the fire is still undetected zero to ten days after it ignited. A similar relationship appears for the lightning-caused models, though the difference in probabilities is smaller. These survival functions make complete sense in the wildland fire context; human-caused fires happen *near* humans and are often detected very quickly, whereas lightning-caused fires take longer to detect since they often occur in more remote locations that are not surveilled 24/7 or they can be smouldering underground undetected until fire conditions worsen (Wotton and Martell, 2005).

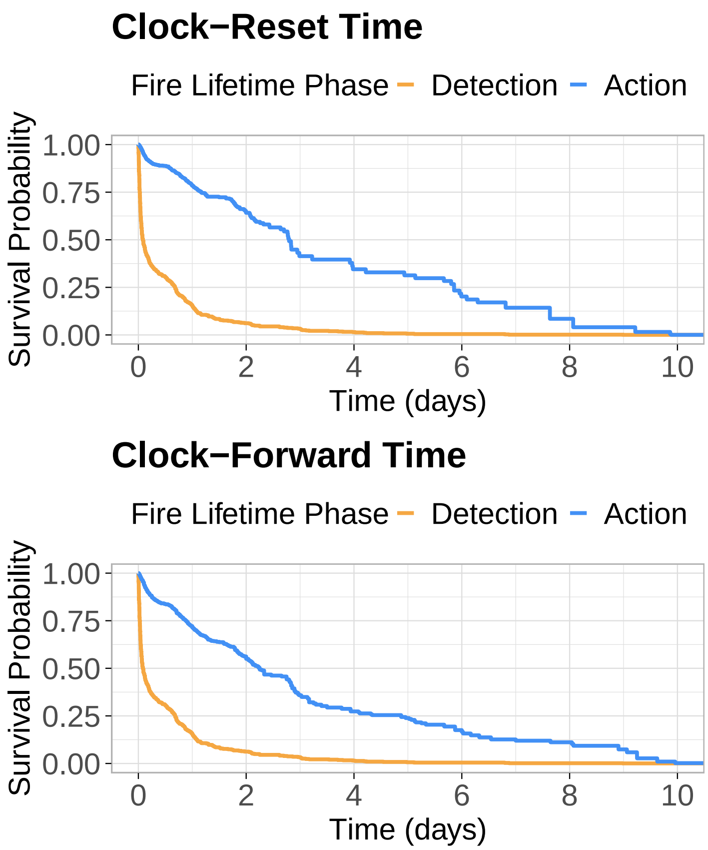


Figure 3.12: Plots of the survival functions for the human-caused clock-reset and clock-forward multi-state models.

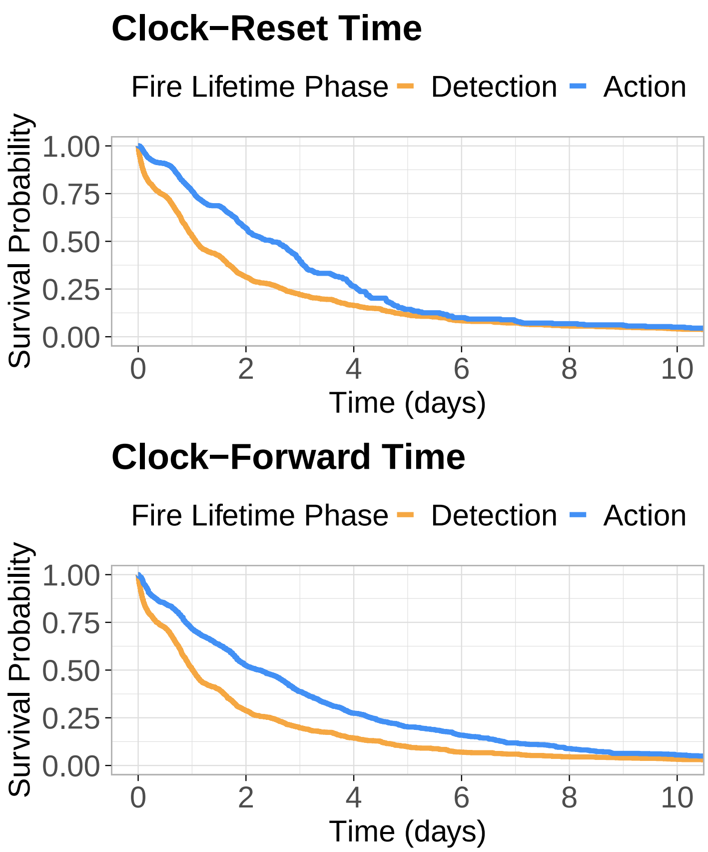


Figure 3.13: Plots of the survival functions for the lightning-caused clock-reset and clock-forward multi-state models.

3.4.1 Model Diagnostics

When fitting a multi-state model where the predictors act multiplicatively on the intensity we make the assumption that we have proportional intensities (i.e., proportional hazards). This assumption is checked using statistical tests and graphical diagnostics based on the scaled Schoenfeld residuals. Grambsch and Therneau (1994) explain the theory behind tests for proportional hazards utilized by the `cox.zph` function from the `survival` package. The proportional hazards tests are essentially tests for nonzero slopes in a generalized linear regression of the rescaled residuals (i.e., the scaled Schoenfeld residuals) on time since the residuals should be independent of time. Therefore finding strong evidence against the null hypothesis of zero slopes, using a 1% significance level, refutes the proportional hazards assumption.

This test was performed for the four fitted models described above. Table 3.7 shows the predictors that violate the proportional hazards assumption for the models. We see that early ignition, same day dispatch, and air tankers consistently violate the proportional hazards assumption in the respective detection and action phase. However, it is not unusual to see violations in the proportional hazards assumption when modelling survival data. Our multi-state models still provide some sense of the predictor effects even if some of the predictors violate the proportional hazards assumption.

Table 3.7: A summary of the predictors that violate the proportional hazards assumption, denoted by an ’X’ mark, across all the fitted multi-state models.

	Human-Caused				Lightning-Caused			
	Clock-Reset		Clock-Forward		Clock-Reset		Clock-Forward	
Predictor	Detection	Action	Detection	Action	Detection	Action	Detection	Action
FFMC					✗	✗	✗	✗
DMC					✗		✗	
FWI								
Mixedwood Fuel		✗				✗		✗
Grass Fuel								
Other Fuel								
Spring								
Fall								
Early Ignition	✗		✗		✗		✗	
Weekday Ignition								
Early Report						✗		✗
Same Day Dispatch		✗		✗		✗		✗
FMH Distance						✗		
AB Distance								
FAB Distance								
Road Distance	✗		✗			✗		
Fire Load								
IA Size								
Ground Forces		✗				✗		✗
Air Tankers		✗		✗		✗		✗
Successful IA						✗		✗
Detection								

Plotting the residuals against time is another option for checking if the proportional hazards assumption is valid; a non-random pattern is evidence that the assumption is violated. Figures 3.14-3.16 show the scaled residuals against some of the predictors mentioned in Table 3.7 for the human-caused clock-reset multi-state model. The solid red line provides a visual tool for determining if the proportional hazards assumption is violated whereby a zero slope indicates no violation and a nonzero slope indicates violation. Although variations from a nonzero slope over time are to be expected, large systematic departures from it are not. For all three plots we see that there is a non-random pattern of the residuals so there is a clear violation of the proportional hazards assumption for these predictors. The rest of the diagnostic plots can be found in Appendix A.

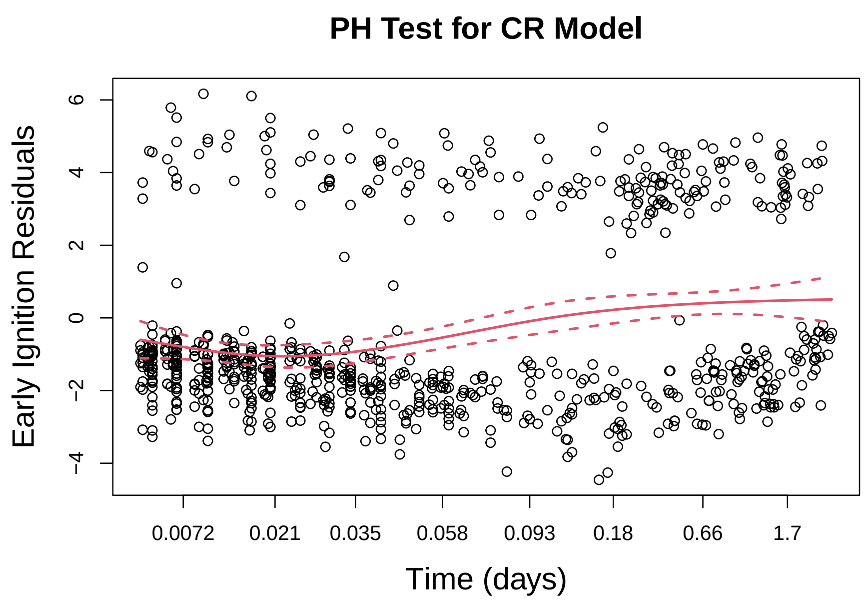


Figure 3.14: Plot of the scaled Schoenfeld residuals for the early ignition predictor from the human-caused CR multi-state model against time in days, along with a smooth curve of the residuals with 95% confidence bands.

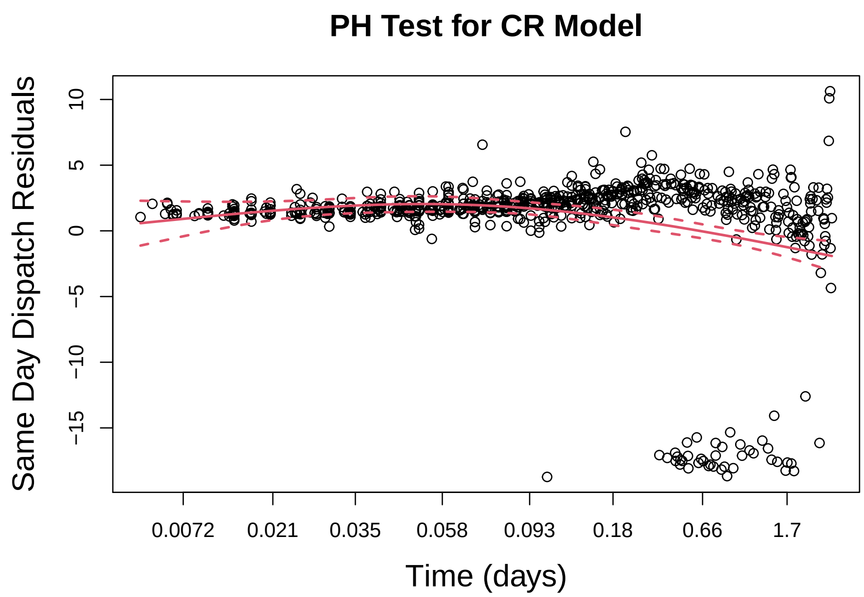


Figure 3.15: Plot of the scaled Schoenfeld residuals for the same day detection predictor from the human-caused CR multi-state model against time in days, along with a smooth curve of the residuals with 95% confidence bands.

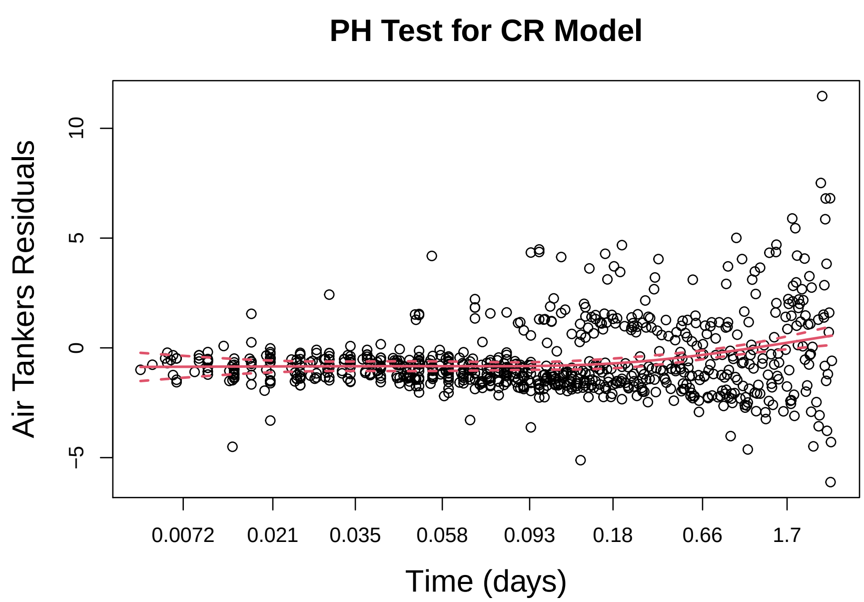


Figure 3.16: Plot of the scaled Schoenfeld residuals for the air tankers predictor from the human-caused CR multi-state model against time in days, along with a smooth curve of the residuals with 95% confidence bands.

3.4.2 Estimated Transition Probabilities

It is often important to obtain prediction probabilities (i.e., estimated transition probabilities on new observations) for multi-state models since we care about knowing how the probability of each possible transition changes over time. We begin by setting up a new dataframe for prediction where each predictor is set to zero and use the `msfit` function from the `mstate` package to calculate the cumulative hazards for our fitted multi-state models on the new data. The resulting incremental outputs over time will be used to calculate the Aalen-Johansen estimator of the transition probabilities, $\widehat{P}_{kl}(0,t)$ (i.e., the probability of being in state l at time t given that you were in state k at time 0).

Figures 3.17 and 3.18 show the estimated transition probabilities for the human-caused clock-reset multi-state model. We see that from the ignition state the probability of staying in ignition decreases to zero and the probability of moving to the report state is quite high for the first day or so but then decreases to zero. This is due to the structure of the 3-progressive process. We also see that from the report state the probability of moving to the ignition state is not possible, the probability of staying in the report state decreases over the first ten days while the probability of moving to the under control state increases over the first ten days. However, we are far less confident in the transition probabilities for transitions into a different state, illustrated by the wide 95% confidence bands in Figure 3.18. Similar conclusions can be drawn from the figures for the other models which can be found in Appendix A.

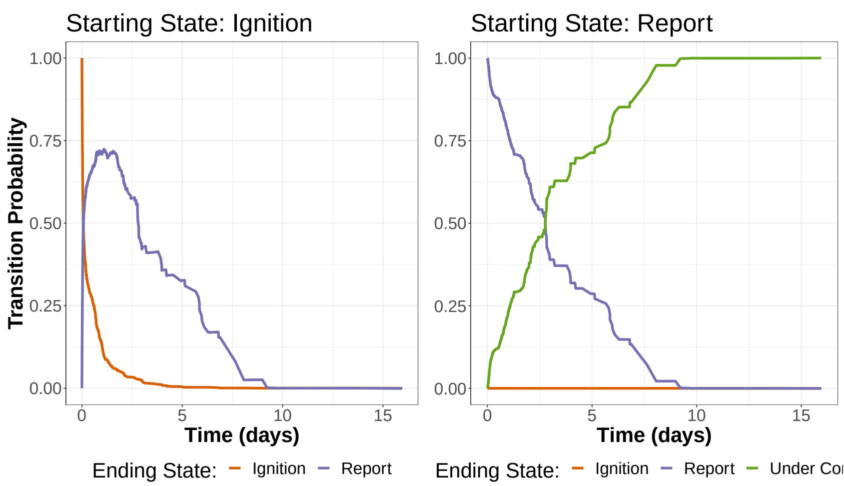


Figure 3.17: Plots of the Aalen-Johansen transition probability estimates for the human-caused clock-reset multi-state model. The left plot shows the $\widehat{P}_{1l}(0,t)$ curves out of the ignition state and the right plot shows the $\widehat{P}_{2l}(0,t)$ curves out of the report state.

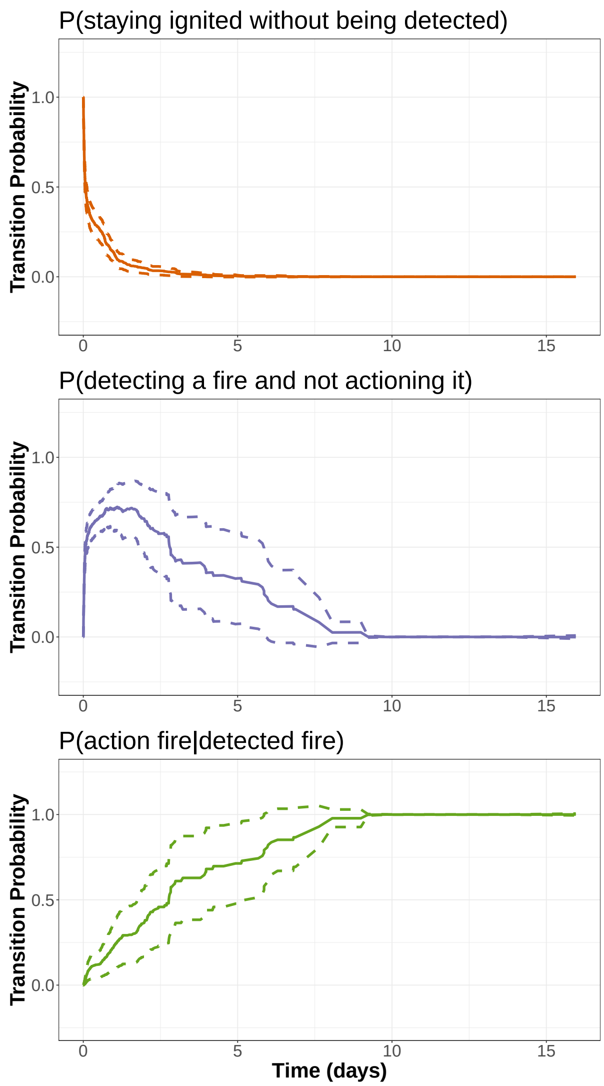


Figure 3.18: Plots of specific Aalen-Johansen transition probabilities taken from Figure 3.17 with the 95% confidence bands.

3.5 Discussion

We found that the clock-reset and clock-forward multi-state models differ from one another with regards to the length of the detection phase being included as a predictor for the action phase. However, there is general consistency in the results from the clock-reset and clock-forward multi-state models, namely corresponding interpretations of model predictors and similar patterns of estimated transition probabilities. Yet the understanding of the fire lifetime phase lengths are vastly different based on the clock times. For clock-reset models time refers to the time since entering a given state, whereas for clock-forward models time is measured since ignition time (Putter et al., 2007). Models that utilize the clock-reset time are preferred because they: (1) provide more information about the two phases due to the two distinct partitions of the fire lifetime, and (2) allow us to fit broader classes of models.

It may be possible to resolve the noted predictor violations of the proportional hazards assumption by adding an interaction between the predictor and time, by stratification, or by also fitting broader classes of models. However, there are cautions that should be considered before adopting stratification. Stratification is useful for nuisance variables with few distinct values when the effect of such a variable is not of direct interest (Fox and Weisberg, 2002). Therefore, careful consideration is required when deciding which variables to stratify by.

One limitation of our work is restricting our fire lifetime dataset to fully suppressed fires with positive phase lengths. Although this restriction allows the fires to perfectly transition through the process, as outlined in Figure 3.7, we may be missing interesting fire lifetimes. For example, some fires were observed to have one or both of their phase lengths be zero. More investigations into the historical fire records to discover why such fires are recorded in this manner is warranted for future work. Another limitation is that we cannot easily specify what the baseline hazard functions are when utilizing this modelling framework, foregoing an added source of flexibility in our model structure.

Moreover, there may be differences between the length of the detection and/or the length of the action phase for fires under near-similar scenarios. This variation may be due to intrinsic differences between different fires. Incorporating a fire-specific random effect would allow for more flexibility when modelling fire lifetimes and it is possible to do so for Cox PH models but becomes more difficult when trying to make predictions (i.e., determine the estimated transition probabilities) using a multi-state model.

Lastly, one must never forget about the end-user's role when developing our models. Although multi-state models overall are pretty straight-forward to understand, they use potentially confusing clock times and there are other methods that can be used to jointly model two outcomes. Therefore we turn to the different modelling technique of joint frailty models that utilize the accessible clock-reset time and a Bayesian framework to add more flexibility to our analysis, interpretations, and ideas.

Chapter 4

Linking Two Phases of Wildland Fire Lifetimes with Joint Frailty Models

One approach to modelling the dependence between two or more outcomes, like the detection and action phases of a wildland fire's lifetime, is by fitting a joint outcome model. This approach offers an enhancement over the 3-progressive multi-state model from Chapter 3 to allow for correlation between the detection and action phases. Our joint frailty modelling framework uses individual fire-specific random effects, which are similar to the common cluster-specific random effects used in frailty models, to incorporate a fire's variation that is common to the outcomes (i.e., phases).

Here, we develop joint frailty models using a Bayesian framework for the Sioux Lookout District wildland fires (786 human-caused fires; 1,270 lightning-caused fires) from Chapter 3. Recall that past research only considered a single phase of the fire's lifetime rather than understanding the fire's evolution over several phases. Through this modelling approach we aim to identify the connection between the distributions of time in the detection and action phases of a fire and determine what factors affect the time spent in these phases.

4.1 Methods

4.1.1 Modelling Framework

For fire i, $i = 1, \ldots, n$, let t_{ij} be the fire lifetime phase duration, specifically the *detection* phase when $j = 1$ and the *action* phase when $j = 2$. Our joint model, or joint frailty model, extends the Cox PH model with the hazard function, h_{ij}, taking the following form

$$h_{ij}(t_{ij}|\mathcal{H}(t_{ij}^{-})) = h_{0j}(t_{ij}|\mathcal{H}(t_{ij}^{-}))u_{ij}\exp(\boldsymbol{x}_{ij}^{'}\boldsymbol{\beta}_j), \tag{4.1}$$

where $h_{0j}(\cdot)$ is the unspecified baseline hazard function common to all fires, $\boldsymbol{x}_{ij}$ is the column vector of predictors corresponding to fixed effects for outcome j, $\boldsymbol{\beta}_j$ is the column vector of fixed effects coefficients for outcome j, $\mathcal{H}(t_{ij}^{-})$ is the history for all events over $[0, t_{ij})$, $\boldsymbol{t}_i = (t_{i1}, t_{i2})'$ is the clock-reset times for the two outcomes (or phases) indicating that the joint frailty model is a semi-Markov model, and u_{ij} is the fire-specific frailty term for the two outcomes which links the two phases and will be discussed in detail later. The frailty term is the critical addition to the model that permits correlation in the outcomes and, depending on how it is specified, correlation can be accommodated in a variety of ways. We define $\boldsymbol{t} = (\boldsymbol{t}_1, \ldots, \boldsymbol{t}_n)'$, $\boldsymbol{x}_i = (\boldsymbol{x}_{i1}, \boldsymbol{x}_{i2})'$, and $\boldsymbol{\beta} = (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2)'$.

The hazard function from (4.1) for the joint model of the detection and action phases of fully suppressed fires contains an unspecified baseline hazard function. We consider two forms for the baseline hazard function based on the framework developed in Nathoo and Dean (2008): a parametric Weibull (W) baseline and a semiparametric piecewise exponential (PE) baseline. A parametric Weibull baseline hazard function has the form $h_{0j}(t_{ij}|\mathcal{H}(t_{ij}^{-})) = \lambda_j \rho_j t_{ij}^{\rho_j - 1}$, where $\lambda_j > 0$ is the scale parameter and $\rho_j > 0$ is the shape parameter. Then (4.1) becomes

$$h_{ij}(t_{ij}|\mathcal{H}(t_{ij}^{-})) = \lambda_j \rho_j t_{ij}^{\rho_j - 1} u_{ij}\exp(\boldsymbol{x}_{ij}^{'}\boldsymbol{\beta}_j), \tag{4.2}$$

where $T_{ij} \sim \text{Weibull}\left(\lambda_j u_{ij} \exp(\boldsymbol{x}_{ij}'\boldsymbol{\beta}_j), \rho_j\right)$, as shown in Appendix B. Thus, the lifetime phase duration of all the fires for the same outcome (detection or action duration) are assumed to follow a Weibull distribution with the same shape parameter ρ_j (ρ_1 and ρ_2, for the detection and action phases, respectively); however, they differ with respect to the scale parameter $\lambda_j u_{ij} \exp(\boldsymbol{x}_{ij}'\boldsymbol{\beta}_j)$ (Duchateau and Janssen, 2007) which is modulated by both predictors and random effects.

We let $\beta_{j0} = \log(\lambda_j)$, $j = 1, 2$, $\boldsymbol{\beta}_0 = (\beta_{10}, \beta_{20})'$, $\boldsymbol{\rho} = (\rho_1, \rho_2)'$, $\boldsymbol{\theta} = (\boldsymbol{\beta}, \boldsymbol{\beta}_0, \boldsymbol{\rho})'$, $\boldsymbol{u}_i = (u_{i1}, u_{i2})'$, and $\boldsymbol{u} = (\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n)'$. Then the joint posterior distribution corresponding to (4.2) is

$$p(\boldsymbol{\theta}, \boldsymbol{u}|\boldsymbol{t}) \propto p(\boldsymbol{t}|\boldsymbol{u}, \boldsymbol{\theta})p(\boldsymbol{u})p(\boldsymbol{\beta})p(\boldsymbol{\beta}_0)p(\boldsymbol{\rho}).$$

The likelihood function becomes

$$p(\boldsymbol{t}|\boldsymbol{u}, \boldsymbol{\theta}) \propto \prod_{i=1}^{n} f(\boldsymbol{t}_i|\boldsymbol{x}_i, \boldsymbol{\theta}, \boldsymbol{u}_i),$$

where $f(\boldsymbol{t}_i|\boldsymbol{x}_i, \boldsymbol{\theta}, \boldsymbol{u}_i)$ is the conditional joint Weibull density function of the outcomes given $\boldsymbol{u}_i$. We assume that the outcomes (i.e., the detection and action durations) are independent given their frailties which results in

$$f(\boldsymbol{t}_i|\boldsymbol{x}_i, \boldsymbol{\theta}, \boldsymbol{u}_i) = \prod_{j=1}^{2} f_j(t_{ij}|\boldsymbol{x}_{ij}, \boldsymbol{\beta}_j, \beta_{0j}, \rho_j, u_{ij}),$$

where $f_j(t_{ij}|\boldsymbol{x}_{ij}, \boldsymbol{\beta}_j, \beta_{0j}, \rho_j, u_{ij})$ is the conditional Weibull density function of outcome j given u_{ij}.

A semiparametric piecewise exponential baseline hazard function has the form

$$h_{0j}(t_{ij}|\mathcal{H}(t_{ij}^-)) = \sum_{k_j=1}^{K_j} \lambda_{jk_j} I_{(a_{k_j-1},\ a_{k_j}]}(t_{ij}),\ j = 1, 2,$$

where K_j denotes the number of intervals for outcome j, a_{k_j} are the join points for outcome j indexed by k_j, and the indicator function $I_{(a_{k_j-1},\ a_{k_j}]}(\cdot)$ is defined as in (2.2).

In this case, (4.1) becomes

$$h_{ij}(t_{ij}|\mathcal{H}(t_{ij}^{-})) = \sum_{k_j=1}^{K_j} \lambda_{jk_j} I_{(a_{k_j-1},\ a_{k_j}]}(t_{ij}) u_{ij} \exp(\boldsymbol{x}_{ij}^{'}\boldsymbol{\beta}_j), \tag{4.3}$$

where $\boldsymbol{\lambda}_j = (\lambda_{j1}, \ldots, \lambda_{jK_j})' > 0$ for all j. For both of the fire lifetime phases we split the time axis into four intervals (i.e., $K_1 = K_2 = 4$) with three join points by placing them near the first, second, and third quartiles. We added 10^{-3} to each of the quartile locations for the join point placements to ensure that they do not perfectly coincide with a fire data point. Alternative partitions may be employed, as appropriate to the context and modelling needs. Kalbfleisch and Prentice (1973) suggested choosing intervals independently of the data but noted that intervals defined by the observed event times resulted in similar hazard curves. In this case, we let $\boldsymbol{\lambda} = (\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2)'$ and $\boldsymbol{\theta} = (\boldsymbol{\beta}, \boldsymbol{\lambda})$. Then the joint posterior distribution for (4.3) is

$$p(\boldsymbol{\theta}, \boldsymbol{u}|\boldsymbol{t}) \propto p(\boldsymbol{t}|\boldsymbol{u}, \boldsymbol{\theta}) p(\boldsymbol{u}) p(\boldsymbol{\beta}) p(\boldsymbol{\lambda}).$$

4.1.2 Fire-Specific Frailty Term

We compare several specifications of the frailty term in the models, as outlined in Table 4.1. Note that $b_i \sim \mathrm{N}(0, \sigma_b^2)$, $d_i \sim \mathrm{N}(0, \sigma_d^2)$, $u_i \sim \mathrm{G}(\psi_u, \psi_u)$, and $v_i \sim \mathrm{G}(\psi_v, \psi_v)$, where $\mathrm{N}(\cdot,\cdot)$ and $\mathrm{G}(\cdot,\cdot)$ are defined in Chapter 2.6.1. Both the Lognormal distribution and the Gamma distribution are common choices in frailty modelling as discussed in Chapter 2.6.1. The one-parameter Gamma distribution with mean equal to one is often used for identifiability purposes (Duchateau and Janssen, 2007; Alvares et al., 2021) when the model includes a general mean term — $\boldsymbol{\beta}_0$ for the Weibull and $\boldsymbol{\lambda}_j$ for the piecewise exponential. The variances of u_i and v_i are $1/\psi_u$ and $1/\psi_v$, respectively. Therefore, we define $\phi_u = 1/\psi_u$ and $\phi_v = 1/\psi_v$ as the parameters; ϕ_u and ϕ_v provide information on the variability (i.e., the heterogeneity) in the population of fires for each respective phase

Table 4.1: Parameterization of the frailty term in the joint model.

Frailty Form	Frailty Distribution	Frailty Term; $\boldsymbol{u}_i = \begin{pmatrix} u_{i1} \\ u_{i2} \end{pmatrix}$
Separate (S)	N/A	$\begin{pmatrix} 1 \\ 1 \end{pmatrix}$
Factor Loading (FL)	Lognormal	$\begin{pmatrix} \exp(b_i) \\ \exp(\gamma b_i) \end{pmatrix}$
	Gamma	$\begin{pmatrix} u_i \\ \gamma u_i \end{pmatrix}$
Factor Loading & Independent (FLI)	Lognormal	$\begin{pmatrix} \exp(b_i) \\ \exp(\gamma b_i + d_i) \end{pmatrix}$
	Gamma	$\begin{pmatrix} u_i \\ \gamma u_i + v_i \end{pmatrix}$
Independent (I)	Lognormal	$\begin{pmatrix} \exp(b_i) \\ \exp(d_i) \end{pmatrix}$
	Gamma	$\begin{pmatrix} u_i \\ v_i \end{pmatrix}$

and will be important parameters to consider in any analysis.

The *separate* form (S) of the joint frailty distribution assumes that there is no fire-specific random effect in the two fire lifetime phases and results in separate Cox PH models for each phase. The *factor loading* form (FL) of the joint frailty distribution uses a factor loading framework on the fire-specific frailty (i.e., $\exp(b_i)$ or u_i, depending on the frailty distribution) between the two phases where the γ parameter accommodates different scales for the effect of the frailty term on the two outcomes. The *factor loading and independent* form (FLI) of the joint frailty distribution is an extension of the FL form whereby an independent frailty is included for each outcome (i.e., d_i or v_i) in addition to the shared factor loading frailty. The addition of this independent frailty term in the model specification of the action phase allows for a fire-specific random effect of that lifetime phase that is not linked with the previous detection phase. Finally, the *independent* form (I) assumes that both the detection and action phases have different fire-specific random effects that are not linked together.

The following interpretations hold for the Lognormal frailties. If $\log(u_{i1}) = b_i$ is greater (less) than zero then the hazard ratio (HR) $u_{i1} = \exp(b_i)$ is greater (less) than one, which results in fire i having a shorter (longer) detection phase after accounting for all other model effects. Similarly if $\log(u_{i2})$ is greater (less) than zero then the HR u_{i2} is greater (less) than one, which results in fire i having a shorter (longer) action phase after accounting for all other model effects. To assist with interpretation, we re-arrange the Gamma frailties from (4.1) in the following way

$$\begin{aligned} h_{ij}(t_{ij}|\mathcal{H}(t_{ij}^-)) &= h_{0j}(t_{ij}|\mathcal{H}(t_{ij}^-))u_{ij}\exp(\boldsymbol{x}_{ij}^{'}\boldsymbol{\beta}_j) \\ &= h_{0j}(t_{ij}|\mathcal{H}(t_{ij}^-))\exp(\boldsymbol{x}_{ij}^{'}\boldsymbol{\beta}_j + \log(u_{ij})), \end{aligned}$$

where $u_{ij} > 0$. In this case, if $u_{i1} = u_i$ is less (greater) than one then $\log(u_i)$ is less (greater) than zero and the HR, $\exp(\log(u_i))$, is less (greater) than one, which results in fire i having a longer (shorter) detection phase after accounting for all other model effects. If u_{i2} is less (greater) than one then $\log(u_{i2})$ is less (greater) than zero and the HR, $\exp(\log(u_{i2}))$, is less (greater) than one, which results in fire i having a longer (shorter) action phase after accounting for all other model effects. The two important aspects of the factor loading parameter, γ, are its sign and whether it has a strong contribution to the model (i.e., if an FL or FLI model form is chosen rather than an S or I form).

4.1.3 Prior and Hyperprior Distributions

For the models with a Weibull baseline hazard function, informative prior distributions were chosen by utilizing prior predictive checks which generate data according to the prior to assess whether the prior is appropriate (Gabry et al., 2019). We simulated parameters according to several different prior distributions, and then simulated fire lifetime phases using the simulated parameters and mean predictor values. Visualizations of the simulated fire phases were created to assess the priors. The following informative prior distributions are used for these models: $\beta_{10}, \beta_{20} \sim \mathrm{N}(0, 0.1^2)$, $\boldsymbol{\beta}_1 \sim \mathrm{N}(0, 0.1^2)$,

$$\boldsymbol{\beta}_2 \sim \begin{cases} \mathrm{N}(0, 0.1^2), & \text{for FL Lognormal frailties when fires are lightning-caused,} \\ \mathrm{N}(0, 0.5^2), & \text{otherwise,} \end{cases}$$

$$\rho_1, \rho_2 \sim \begin{cases} \text{half-N}(0, 3^2), & \text{for FL Lognormal frailties when fires are lightning-caused,} \\ \text{half-N}(0, 2^2), & \text{otherwise.} \end{cases}$$

When the factor loading variable γ is present in the model, we use the following prior distribution

$$\gamma \sim \begin{cases} \mathrm{N}(0, 0.1^2), & \text{for Lognormal frailties,} \\ \text{half-N}(0, 0.01^2), & \text{for Gamma frailties.} \end{cases}$$

Also, the following hyperprior distributions are used: $\psi_u, \psi_v \sim \mathrm{G}(0.5, 0.5)$, $\omega_b = 1/\sigma_b^2 \sim \mathrm{G}(4, 2)$, and $\omega_d = 1/\sigma_d^2 \sim \mathrm{G}(4, 2)$. The inverse-Gamma hyperprior distributions for the variance parameters of the Lognormal fire-specific random effects were chosen based on the recommendations in Gelman (2006) and Korsgaard et al. (1998).

For the models with a piecewise exponential baseline hazard function, certain prior and hyperprior distributions stay the same (i.e., $\boldsymbol{\beta}_1, u_i, v_i, b_i, d_i, \gamma$) while the other distributions are: $\boldsymbol{\beta}_2 \sim \mathrm{N}(0, 0.5^2)$, $\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2, \psi_u, \psi_v \sim \mathrm{G}(0.01, 0.01)$, and

$$\omega_b, \omega_d \sim \begin{cases} \mathrm{G}(4, 4), & \text{for lightning-caused fires,} \\ \mathrm{G}(4, 2), & \text{for human-caused fires.} \end{cases}$$

4.1.4 Bayesian Modelling

We model human- and lightning-caused fires separately. For each case, we fit joint models using the Weibull and piecewise exponential baselines. The combinations of baseline hazard functions, frailty forms, and frailty distributions previously discussed result in 14 different models. We consider and contrast all of these, and use the same predictors for these models as those from the multi-state models, shown in Table 3.3. Bayesian techniques are used to fit the models to both the human- and lightning-caused wildland fire data, using JAGS. See Chapter 2.7 for more details. Model fitting is carried out by adaptive MCMC using the R package `runjags` (Denwood, 2016) with three chains. Each chain has 30,000 adaptive steps, 30,000 burn-in steps, and 12,000 steps where samples are thinned at every fourth step to reduce autocorrelation. Chains are run on parallel hardware to improve computational efficiency.

The parameter estimates presented in Section 4.2 are the posterior means; additionally, the posterior medians (50% quantiles) are also provided. Convergence is assessed by visually examining the chain trajectories and density plots of the sampled parameter values along with calculating the Gelman-Rubin statistic, or the potential scale reduction statistic, $\widehat{R}$ (Gelman and Rubin, 1992). The credible intervals are obtained as the lower/upper 5% quantiles of the posterior density. The number of effective samples provide a measure of how much independent information there is in the autocorrelated chains (Kruschke, 2015). As the likelihood function for any model using the piecewise exponential baseline hazard function shown in (4.3) is not implemented in JAGS, we use the "Poisson-zeros" method outlined in Ntzoufras (2009) to specify it directly. Model comparisons are performed for the human- and lightning-caused models using both the WAIC and Pareto smoothed importance sampling leave-one-out (PSIS-LOO) methods discussed in Chapter 2.8.

4.1.5 Visualizations

Scatterplot visualizations in Section 4.2 utilize smooth local regression lines, calculated using the locally estimated scatterplot smoothing (LOESS) method. The nonparametric local regression uses the data from a neighbourhood around the specific data point to generate a weighted least-squares estimate. In our case, the neighbourhood is defined by a span of 0.75, indicating that the closest three-quarters of the total data points are used as the neighbourhood. For more details on local regression see Montgomery et al. (2012).

4.2 Analysis and Results

4.2.1 Human-Caused Fires

Comparisons of the expected log pointwise predictive densities (ELPD) using WAIC and PSIS-LOO for human-caused fires are shown in Figure 4.1. The preferred model with the largest ELPD (Vehtari et al., 2020) is highlighted in red. Model ranking using the two methods does not perfectly correspond, but the preferred model under both criteria is shown to be of FLI form with Gamma frailties (see Table 4.1) using a Weibull baseline for human-caused fires. Figures showing these ranked ELPD estimates can be found in Appendix B. Although not shown here, the estimates of the effective number of parameters corresponding to the two methods, WAIC and LOOIC[1], have a range of (26, 33) for models without frailty terms and a higher range of (218, 529) for models with frailty terms since we are estimating individual fire-specific random effects.

A summary of the posterior estimates for this model is shown in Table 4.2. Recall that HR estimates, here calculated using the mean posterior estimates, that are greater (less) than one imply a shorter (longer) phase length with all other predictors being held equal.

[1]LOOIC stands for the **leave-one-out information criterion** that is calculated using the PSIS-LOO method (i.e., $\widehat{\text{LOOIC}} = -2 \times \widehat{\text{ELPD}}_{\text{PSIS-LOO}}$).

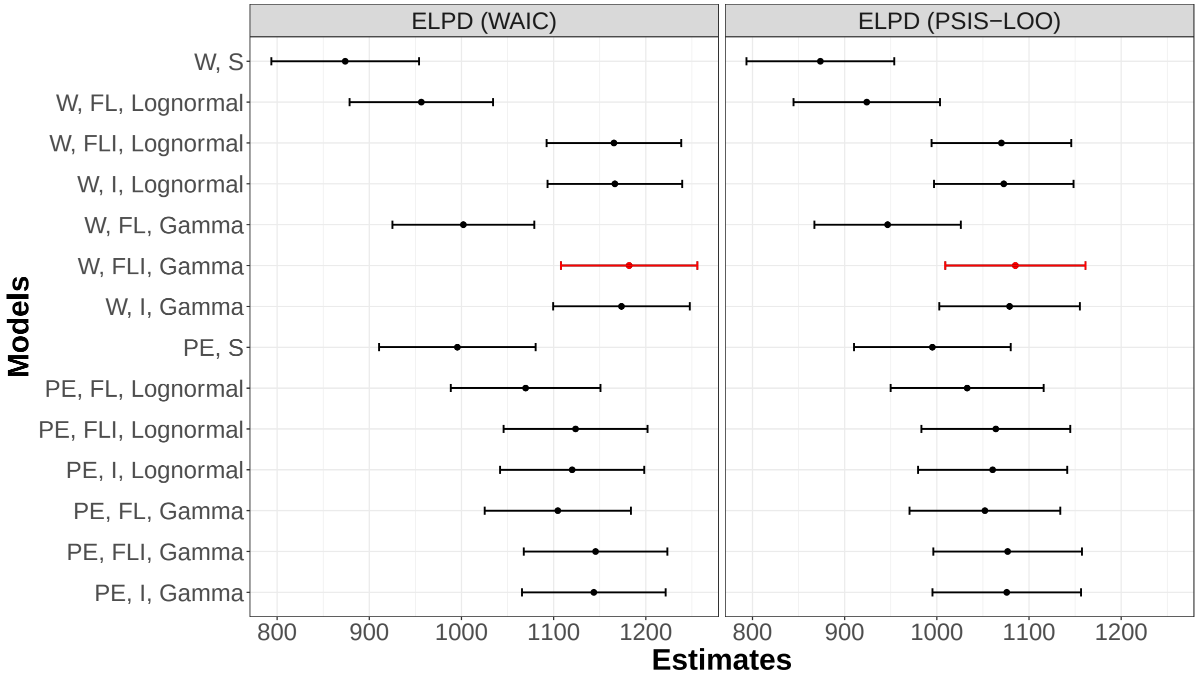


Figure 4.1: Comparisons of ELPD point estimates and standard errors, using the WAIC and PSIS-LOO methods, for human-caused wildland fires. The preferred model is highlighted in red.

Although not shown here, the HR estimates for the predictor coefficients are generally consistent across the 14 models in terms of being greater or less than one, except for the predictors with no evidence of having an effect on their respective fire lifetime phases; however, their credible intervals differ across models. All parameter estimates reached convergence with $\widehat{R} \leq 1.01$. Refer to Table 3.2 for descriptions of these variables. The interpretations of the predictors whose credible intervals do not overlap zero from this **human-caused** joint frailty model follow:

- **FWI:** Fires burning under more intense conditions are associated with a shorter detection phase.

- **Fuel:** Fires burning from grass fuel have shorter detection and action phases than those burning from coniferous fuel.

- **Season:** Fires ignited in spring have shorter detection phases than those in summer, whereas fires ignited in fall have longer detection and action phases than those in summer.

- **Time of Ignition:** Fires ignited earlier in the day have a longer detection phase than those ignited after noon while fires ignited during a weekday have a shorter detection phase than those ignited on the weekend.

- **Same Day Dispatch:** Fires that have same day dispatch tend to have a much shorter action phase than those that do not.

- **Distance:** Human-caused fires located farther away from the AB or roads are associated with a longer detection phase. However, human-caused fires located farther away from the FAB are associated with a shorter detection phase. Similarly, human-caused fires located farther away from the FMH or AB are associated with a longer action phase.

- **Suppression Efforts:** Fires that require more suppression efforts are associated with a longer action phase.

- **Successful IA:** Fires that have a successful initial attack have a much shorter action phase than those that do not.

- **Detection:** Fires with longer detection phases are associated with longer action phases.

We notice that while we have similar interpretations of the predictor effects as from the previous chapter, using this joint modelling framework we are able to identify that a longer detection phase is associated with a longer action phase.

Table 4.2: Summary of posterior estimates for the human-caused W, FLI, Gamma model.

Parameters	Mean	HR Estimate	5% Quantile	50% Quantile	95% Quantile	Number of Effective Samples
Coefficients for Detection Predictors						
FWI	0.10 (0.04)	1.11	0.04	0.11	0.17	31,893
Mixedwood Fuel	−0.05 (0.08)	0.96	−0.17	−0.04	0.08	31,938
Grass Fuel	0.34 (0.07)	1.41	0.22	0.34	0.47	31,716
Other Fuel	0.09 (0.08)	1.10	−0.03	0.09	0.22	33,133
Spring	0.45 (0.07)	1.56	0.33	0.45	0.56	19,026
Fall	−0.24 (0.08)	0.78	−0.38	−0.24	−0.11	31,087
Early Ignition	−0.19 (0.07)	0.83	−0.30	−0.19	−0.07	31,527
Weekday Ignition	0.30 (0.07)	1.36	0.19	0.30	0.42	12,228
AB Distance	−0.17 (0.04)	0.85	−0.23	−0.17	−0.10	30,014
FAB Distance	0.22 (0.04)	1.24	0.15	0.22	0.28	18,304
Road Distance	−0.32 (0.07)	0.73	−0.43	−0.32	−0.21	26,079
Coefficients for Action Predictors						
Mixedwood Fuel	−0.24 (0.19)	0.78	−0.56	−0.24	0.06	8,929
Grass Fuel	0.72 (0.17)	2.06	0.45	0.72	1.00	8,017
Other Fuel	0.09 (0.17)	1.09	−0.20	0.09	0.37	9,666
Spring	0.13 (0.14)	1.13	−0.11	0.12	0.37	6,320
Fall	−0.52 (0.21)	0.60	−0.87	−0.52	−0.17	9,363
Same Day Dispatch	2.52 (0.26)	12.37	2.09	2.51	2.96	1,031
FMH Distance	−0.23 (0.06)	0.79	−0.34	−0.23	−0.13	10,119
AB Distance	−0.19 (0.06)	0.82	−0.29	−0.19	−0.10	10,123
Ground Forces	−0.06 (0.02)	0.94	−0.09	−0.06	−0.04	5,903
Air Tankers	−1.34 (0.12)	0.26	−1.54	−1.34	−1.15	2,770
Successful IA	1.57 (0.24)	4.81	1.19	1.57	1.97	1,336
Detection	−0.13 (0.07)	0.88	−0.24	−0.13	−0.02	10,679

Parameters	Mean	HR Estimate	5% Quantile	50% Quantile	95% Quantile	Number of Effective Samples
Other Model Parameters						
ρ_1	0.62 (0.02)		0.58	0.62	0.65	$2,499$
ρ_2	1.61 (0.08)		1.47	1.60	1.75	938
β_{10}	0.66 (0.07)		0.54	0.66	0.78	$5,834$
β_{20}	0.01 (0.10)		-0.15	0.01	0.17	$7,234$
ψ_u	6.61 (2.07)		4.12	6.19	10.53	950
ψ_v	0.86 (0.11)		0.69	0.85	1.06	$1,074$
γ	0.01 (0.004)		0.002	0.01	0.01	$9,939$

The variance estimates for the detection and action frailties for this preferred model are 0.16 (0.09, 0.24) and 1.19 (0.94, 1.46), respectively, identifying that there is more variability in the action phase frailties. Figure 4.2 visualizes this result with a spatial plot of the posterior estimates of the frailties for each fire by phase in the Sioux Lookout District. Figure 4.3 provides a temporal plot of the posterior estimates of the frailties for each phase by year. For Figures 4.2 and 4.3, deeper light greens imply a shorter phase length whereas deeper pinks imply a longer phase length. An interesting "corridor" of human-caused fires that took longer to action exists between the Pickle Lake Attack Base and the Savant Lake Forward Attack Base. Figure 4.3 shows that there is no obvious temporal trend for the detection phase but that the action phase length may be getting longer over time. To further investigate these trends we plot the median values from the detection and action densities in Figure 4.3 against year in Figure 4.4. We see that the action phase for human-caused fires in Sioux Lookout appears to be getting longer over time whereas the detection phase is generally consistent.

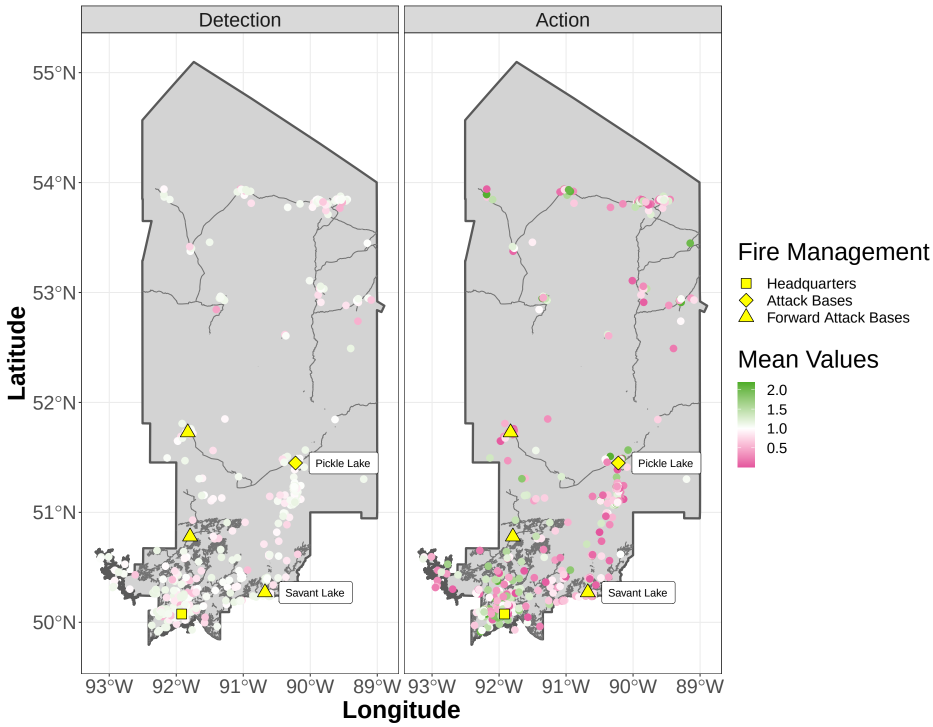


Figure 4.2: Spatial plot of the human-caused posterior estimates for the detection and action frailties. The locations of the fire management headquarters, attack bases and forward attack bases are highlighted in yellow.

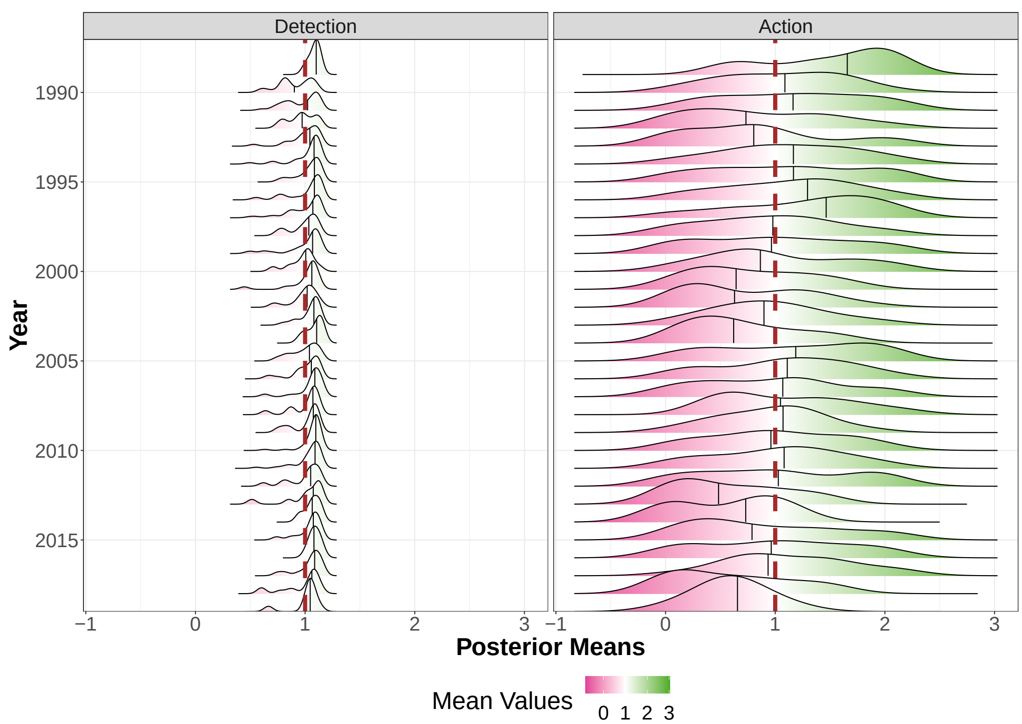


Figure 4.3: Temporal plot of the human-caused posterior estimates for the detection and action frailties over the years. Black solid vertical lines are the medians of the respective densities. Brown dashed lines are the thresholds where the frailty interpretation changes.

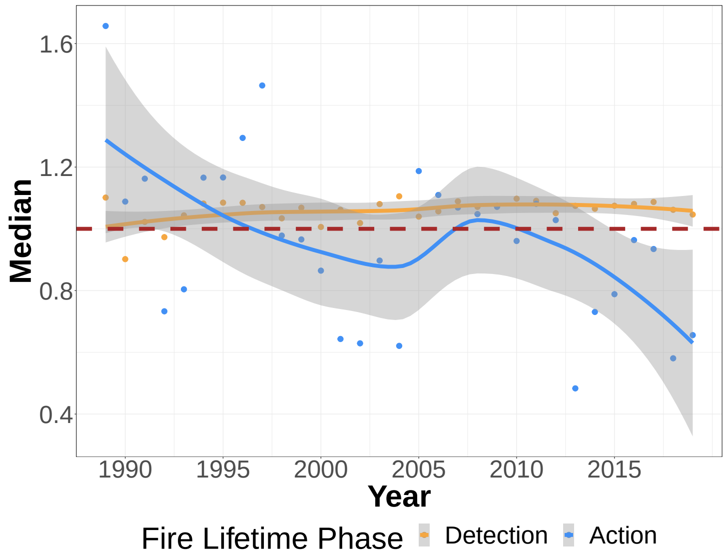


Figure 4.4: Scatterplot of the median values from the densities of the posterior estimates for the human-caused frailties shown in Figure 4.3. Brown dashed line is the threshold where the frailty interpretation changes.

4.2.2 Lightning-Caused Fires

For lightning-caused fires, Figure 4.5 provides the comparisons of the fit of the models. The preferred model with the largest ELPD is highlighted in red; for lightning-caused fires it is of FLI form with Lognormal frailties using a piecewise exponential baseline. Figures of the ranked ELPD estimates can be found in Appendix B. Again, models with frailty terms have a higher range of estimates of the effective number of parameters than models that do not.

Table 4.3 gives a summary of the posterior estimates for this model. Again, we note for completeness that the HR estimates for the predictor coefficients are generally consistent across the 14 models in terms of being greater or less than one, except for the predictors with no evidence of having an effect on their respective fire lifetime phases; however, their credible intervals still differ across models. All parameter estimates reached convergence with $\widehat{R} \leq 1$. The interpretations of the predictors whose credible intervals do not overlap zero from this **lightning-caused** joint frailty model follow:

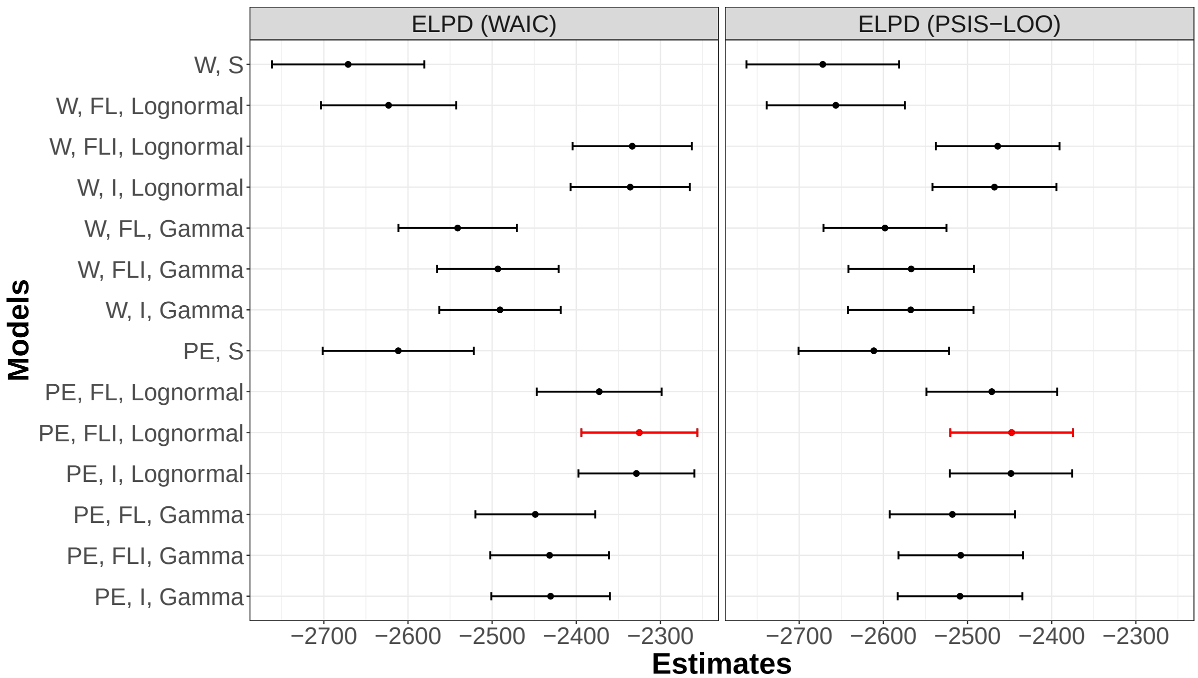


Figure 4.5: Comparisons of ELPD point estimates and standard errors, using the WAIC and PSIS-LOO methods, for lightning-caused wildland fires. The preferred model is highlighted in red.

- **FFMC & DMC:** Fires burning under drier conditions of the smallest forest fuels and medium-sized fuels are associated with a shorter detection phase. Whereas fires burning under drier conditions of only the smallest forest fuels are associated with a longer action phase.

- **Fuel:** Fires burning from mixedwood fuel have a longer detection phase than those burning from coniferous fuel, whereas fires burning from other fuel types have a shorter action phase than those burning from coniferous fuel.

- **Time of Ignition & Report:** Fires ignited earlier in the day have a longer detection phase than those ignited after noon.

- **Same Day Dispatch:** Fires that have same day dispatch tend to have a much shorter action phase than those that do not.

- **Distance:** Lightning-caused fires located farther away from the FMH or roads are associated with a longer action phase.

- **Fire Load & IA Size:** Fires burning when there are more fires already burning on the landscape are associated with a longer action phase. Also, fires with a larger initial attack size are associated with a longer action phase.
- **Suppression Efforts:** Fires that require more suppression efforts are associated with a longer action phase.
- **Successful IA:** Fires that have a successful initial attack have a much shorter action phase than those that do not.
- **Detection:** Fires with longer detection phases are associated with longer action phases.

Table 4.3: Summary of posterior estimates for the lightning-caused PE, FLI, Lognormal model.

Parameters	Mean	HR Estimate	5% Quantile	50% Quantile	95% Quantile	Number of Effective Samples
Coefficients for Detection Predictors						
FFMC	0.19 (0.04)	1.21	0.13	0.19	0.26	11,263
DMC	0.14 (0.04)	1.15	0.07	0.14	0.21	14,909
Mixedwood Fuel	−0.13 (0.08)	0.88	−0.25	−0.13	−0.003	26,276
Grass Fuel	0.001 (0.10)	1.00	−0.16	0.002	0.16	34,555
Other Fuel	−0.0005 (0.08)	1.00	−0.14	−0.0002	0.14	28,343
Early Ignition	−0.28 (0.07)	0.75	−0.39	−0.28	−0.17	18,160
Coefficients for Action Predictors						
FFMC	−0.08 (0.03)	0.92	−0.14	−0.08	−0.03	23,293
Mixedwood Fuel	0.05 (0.10)	1.06	−0.12	0.05	0.22	21,867
Grass Fuel	0.04 (0.33)	1.04	−0.52	0.04	0.58	32,666
Other Fuel	0.46 (0.12)	1.59	0.27	0.46	0.66	26,162
Early Report	0.02 (0.11)	1.02	−0.16	0.02	0.20	24,173
Same Day Dispatch	1.13 (0.11)	3.08	0.94	1.13	1.31	3,615
FMH Distance	−0.28 (0.04)	0.75	−0.35	−0.28	−0.21	19,247
Road Distance	−0.11 (0.03)	0.90	−0.16	−0.11	−0.06	23,182
Fire Load	−0.08 (0.03)	0.92	−0.14	−0.08	−0.03	25,225
IA Size	−0.05 (0.03)	0.95	−0.10	−0.05	−0.003	24,089
Ground Forces	−0.07 (0.01)	0.93	−0.09	−0.07	−0.05	5,400
Air Tankers	−0.55 (0.05)	0.58	−0.63	−0.54	−0.46	8,590
Successful IA	1.43 (0.15)	4.17	1.18	1.43	1.68	1,704
Detection	−0.05 (0.02)	0.95	−0.08	−0.05	−0.02	7,694

Parameters	Mean	HR Estimate	5% Quantile	50% Quantile	95% Quantile	Number of Effective Samples
Other Model Parameters						
λ_{11}	1.04 (0.09)		0.89	1.04	1.20	$2,984$
λ_{12}	0.69 (0.05)		0.61	0.69	0.78	$17,320$
λ_{13}	1.26 (0.14)		1.05	1.25	1.50	$1,869$
λ_{14}	1.30 (0.24)		0.95	1.27	1.71	$1,176$
λ_{21}	0.27 (0.06)		0.19	0.27	0.37	$1,476$
λ_{22}	0.20 (0.04)		0.14	0.20	0.27	$1,471$
λ_{23}	0.63 (0.12)		0.45	0.62	0.84	$1,496$
λ_{24}	0.68 (0.13)		0.49	0.66	0.91	$1,747$
σ_b	1.04 (0.12)		0.85	1.04	1.23	973
ω_b	0.97 (0.23)		0.66	0.93	1.40	908
σ_d	0.62 (0.05)		0.54	0.62	0.71	$1,884$
ω_d	2.64 (0.46)		1.96	2.59	3.46	$1,850$
γ	0.08 (0.07)		-0.03	0.08	0.19	$7,170$

There is more variability in the detection phase frailties for this preferred model than in the action phase frailties, as illustrated in Figure 4.6 which plots the posterior estimates of the frailties for each fire by phase in the Sioux Lookout District. Figure 4.7 provides a temporal plot of the posterior estimates of the frailties for each phase by year. For Figures 4.6 and 4.7, deeper dark greens imply shorter phase lengths and deeper purples imply longer phase lengths. Figure 4.7 shows potential trends for both phases over time. Figure 4.8 plots the median values from the densities of the posterior estimates of the frailties over time. The action phase for lightning-caused fires appears to be getting longer over time and the detection phase exhibits no clear trend.

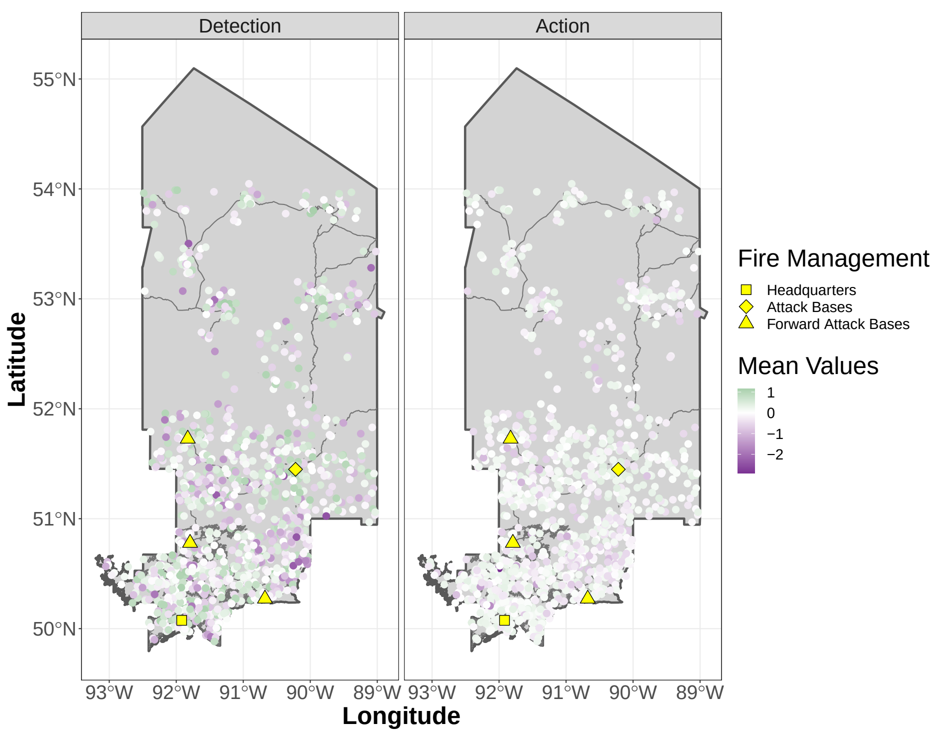


Figure 4.6: Spatial plot of the lightning-caused posterior estimates for the detection and action frailties. The locations of the fire management headquarters, attack bases and forward attack bases are highlighted in yellow.

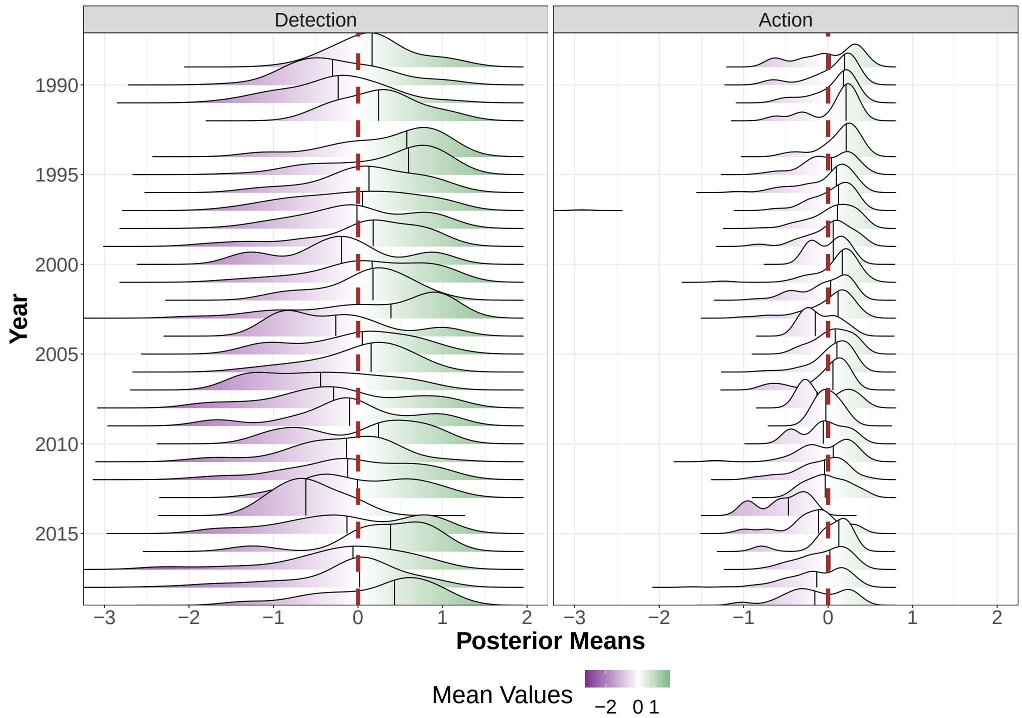


Figure 4.7: Temporal plot of the lightning-caused posterior estimates for the detection and action frailties over the years. Black solid vertical lines are the medians of the respective densities. Brown dashed lines are the thresholds where the frailty interpretation changes.

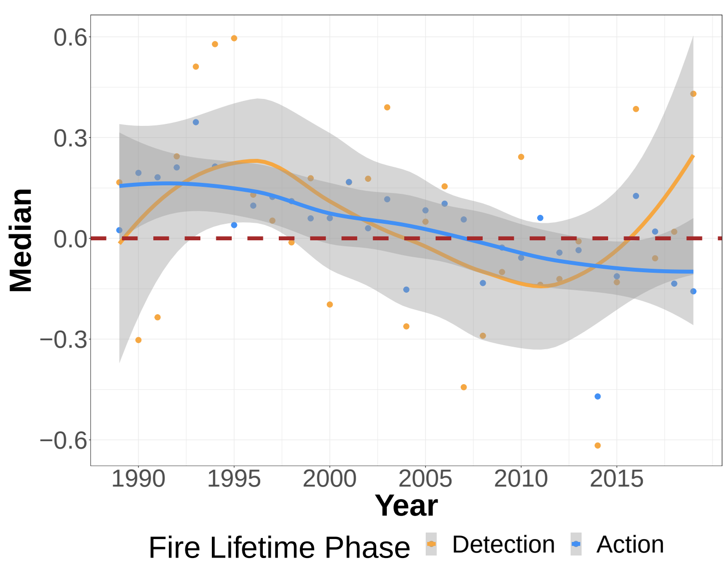


Figure 4.8: Scatterplot of the median values from the densities of the posterior estimates for the lightning-caused frailties shown in Figure 4.7. Brown dashed line is the threshold where the frailty interpretation changes.

4.2.3 Model Diagnostics

The Pareto smoothed importance sampling (PSIS) method for estimating the leave-one-out (LOO) expected log pointwise predictive density, a measure of predictive accuracy, was utilized to compare the models for each cause. Figures 4.9 and 4.10 show the PSIS diagnostic plots for the preferred models discussed in Section 4.2. The Pareto k values, discussed in Chapter 2.8.2, are displayed on the y-axis and the data points (or fires) are shown on the x-axis. Any $\hat{k} > 0.7$ are considered problematic (Vehtari et al., 2017). To rectify problematic Pareto k values they suggest one should: (1) sample directly from $p(\boldsymbol{\theta}^s|\boldsymbol{y}_{-i})$, the posterior evaluated using the $\boldsymbol{\theta}^s$ draws from the full posterior given the data without the ith data point, so long as the number of problematic data points is low; (2) use K-fold cross-validation; or (3) use a more robust model.

The number of problematic fires (i.e., fires with $\hat{k} > 0.7$) shown for the preferred models is quite large making the first option unfeasible. Before attempting the K-fold cross validation we wanted to investigate which fires result in these problematic $\hat{k}$ values across all human- and lightning-caused fire models utilized in the previous section, although we know some models are preferred using previous selection methods. The following analysis of problematic $\hat{k}$ values is conducted for the 14 human-caused joint frailty models discussed earlier. See Appendix B for related diagnostics for the lightning-caused fire models.

We found that 399 (or 51%) of the human-caused fires have $\hat{k}$ values less than 0.7 across all 14 models. These fires had an average detection length of 0.08 days, an average action length of 0.19 days, and an average total duration length of 0.28 days. In contrast, only 1 of the human-caused fires has $\hat{k}$ values greater than or equal to 0.7 across all 14 models. This fire had a detection length of 2.09 days, an action length of 6.81 days, and a total duration length of 8.91 days — a fire that lasted much longer than the 399 fires mentioned. Table 4.4 provides the summary of the number of fires with $\hat{k} \geq 0.7$ for m models, $m = 0, 1, \ldots, 14$, along with some summary statistics. The 399 "unproblematic"

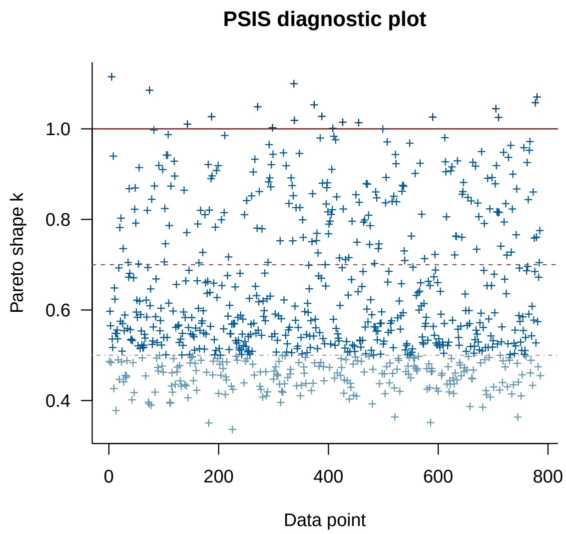


Figure 4.9: Pareto smoothed importance sampling diagnostic plot for preferred human-caused model.

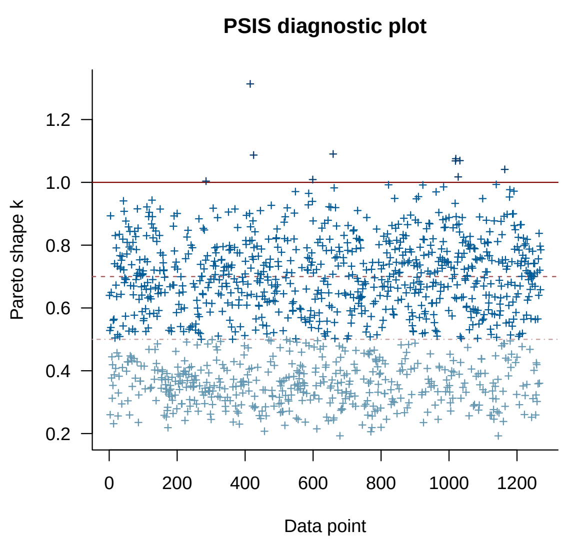


Figure 4.10: Pareto smoothed importance sampling diagnostic plot for preferred lightning-caused model.

fires are not having a strong influence in our models, but several fires do have an influence. For example, 19 (or 2%) of the human-caused fires have $\hat{k}$ values greater than or equal to 0.7 for 12 out of the 14 models, specifically all of the models that include the fire-specific random effects.

Table 4.4: Summary of human-caused fires with $\hat{k} \geq 0.7$. The average total duration length is the sum of the average detection and action lengths, where the lengths are in days.

m	# of Fires with $\hat{k} \geq 0.7$ for m Models	Proportion of Fires with $\hat{k} \geq 0.7$ for m Models	Average Detection Length	Average Action Length	Average Total Duration Length
0	399	0.51	0.08	0.19	0.28
1	68	0.09	0.31	0.21	0.52
2	64	0.08	0.51	0.40	0.91
3	32	0.04	0.54	0.65	1.19
4	45	0.06	0.32	0.93	1.25
5	27	0.03	0.27	1.22	1.49
6	31	0.04	0.57	1.36	1.93
7	14	0.02	0.56	1.18	1.73
8	19	0.02	0.90	0.70	1.60
9	29	0.04	0.35	1.27	1.62
10	20	0.03	1.24	0.90	2.13
11	18	0.02	0.99	1.22	2.21
12	19	0.02	3.16	1.88	5.04
14	1	0	2.09	6.81	8.91

Figures 4.11 and 4.12 illustrate that as m increases, the durations of the fire phases also increase. This increase occurs gradually for the detection phase lengths but is steeper for the action phase lengths. Therefore, we know that fires with longer phase lengths have a stronger influence in the models and result in problematic Pareto k values when the model fit is assessed. This issue was explored temporally across fire years and spatially across fire locations in the Sioux Lookout District - no strong relationships appeared for either. Models with Gamma frailties had fewer $\widehat{k} \geq 0.7$ than the Lognormal frailty models. For models with the FL (FLI or I) form, those that had a Weibull (PE) baseline had fewer $\widehat{k} \geq 0.7$ when compared with the PE (Weibull) baseline.

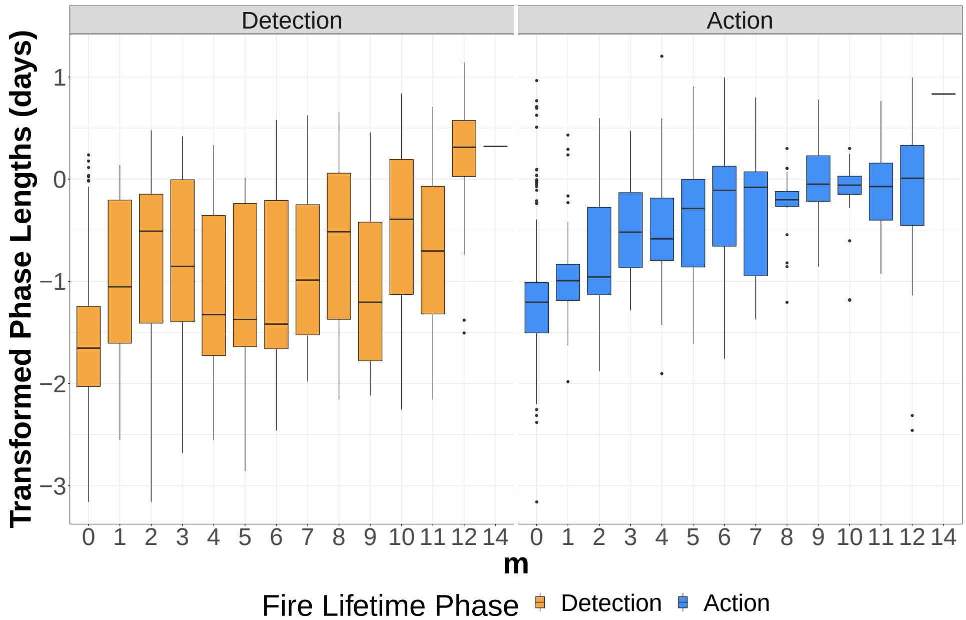


Figure 4.11: Boxplot showing the spread of the log 10 transformed and stratified phase lengths (days) of the associated fires against m from Table 4.4.

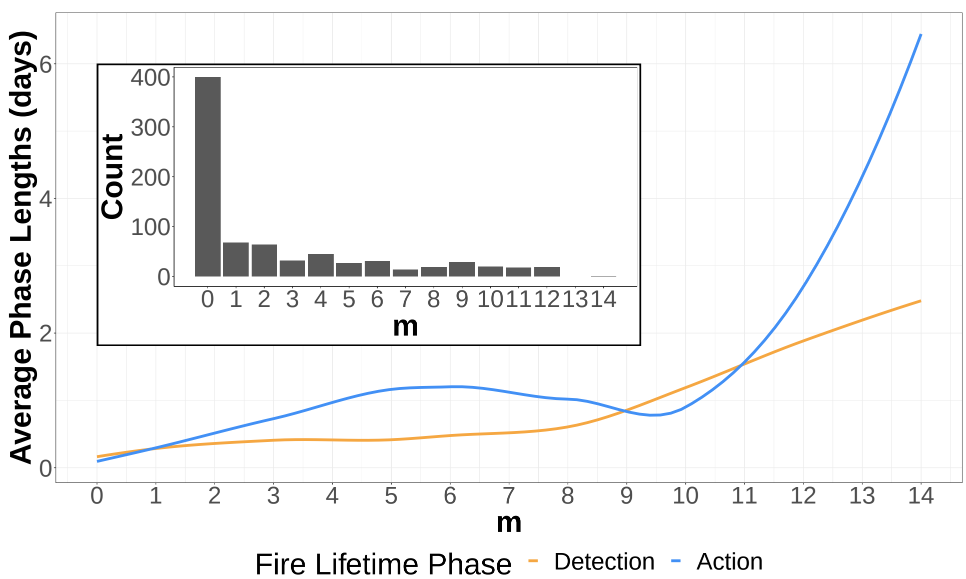


Figure 4.12: Smooth local regression lines of the stratified average phase lengths (days) of the associated fires against m from Table 4.4. A histogram inset is provided to show the fire counts.

To address the concern that fires with longer phase lengths have greater influence in the models, we removed all fires that had $\widehat{k} \geq 0.7$ for at least one of the models (i.e., 51% of the original fires were retained). All 14 models were fit to the 399 fires in JAGS using the same combinations of baselines, frailty types, model forms, and prior distributions discussed previously. Table 4.5 and Figures 4.13 and 4.14 show that the same issue of influential fires occurs even when reducing the data, but that it occurs on a smaller scale (i.e., fewer models have $\widehat{k} \geq 0.7$). We posit that this issue will continually occur due to the structure of the lifetime data — taking out fires with longer lifetime phases will be replaced by fires now considered to have relatively "longer" lifetime phases. Based on the recommendations of Vehtari et al. (2017), either performing a K-fold (e.g., 10-fold) cross validation to calculate the WAIC and PSIS-LOO estimates or fitting a more robust model that allows for added flexibility in the tails when fitting the fire lifetime phase lengths are the only feasible options since the number of problematic data points is large for our human-caused ($\approx$ 49%) and lightning-caused ($\approx$ 63%) fires.

Table 4.5: Summary of human-caused fires with $\hat{k} \geq 0.7$. These models were fit using the unproblematic fires from of the original data. The average total duration length is the sum of the average detection and action lengths, where the lengths are in days.

n	# of Fires with $\hat{k} \geq 0.7$ for n Models	Proportion of Fires with $\hat{k} \geq 0.7$ for n Models	Average Detection Length	Average Action Length	Average Total Duration Length
0	358	0.90	0.04	0.18	0.22
1	17	0.04	0.20	0.37	0.57
2	7	0.02	0.39	0.20	0.59
3	15	0.04	0.85	0.18	1.02
4	1	0	0.26	0.02	0.28
5	1	0	1.30	0.93	2.23

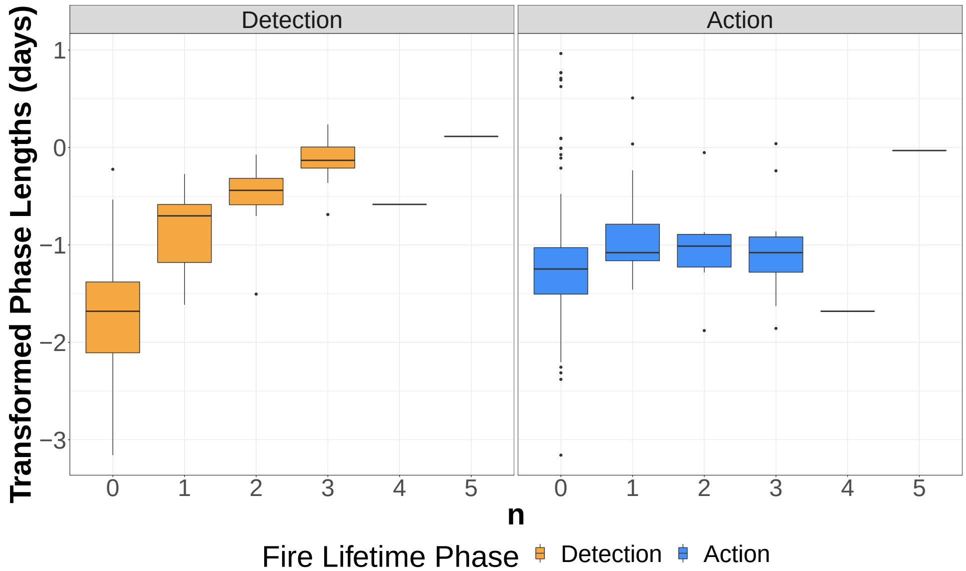


Figure 4.13: Boxplot showing the spread of the log 10 transformed and stratified phase lengths (days) of the associated fires against n from Table 4.5.

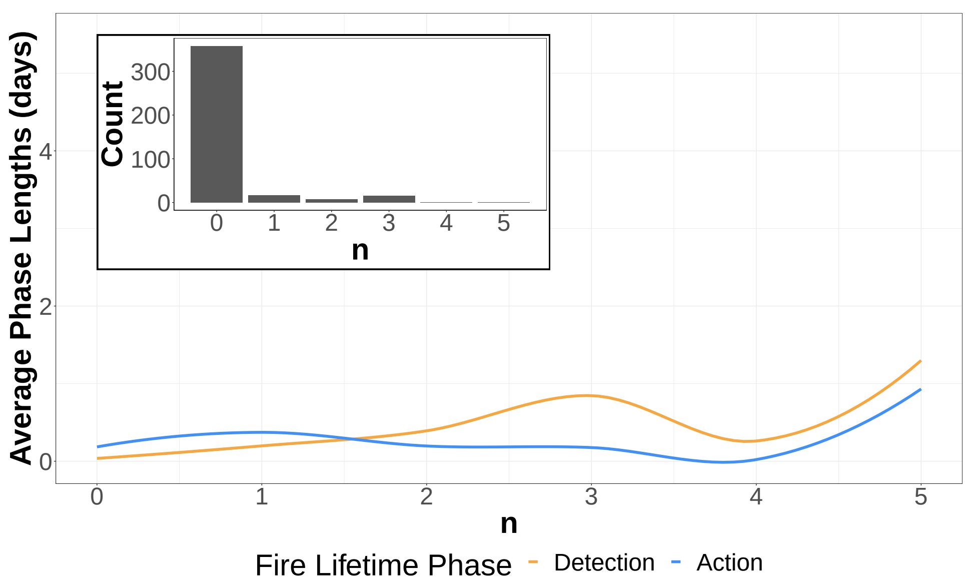


Figure 4.14: Smooth local regression lines of the stratified average phase lengths (days) of the associated fires against n from Table 4.5. A histogram inset is provided to show the fire counts.

4.3 A Simulation Study to Compare Frailty Forms

In Figures 4.1 and 4.5 we see that the FLI and I forms dominate the higher ranked models using both the WAIC and PSIS-LOO goodness-of-fit methods for human- and lightning-caused wildland fires. In fact, the FLI form ranks first across all the cases that we examined. But how often would concluding that the independent random effects are needed in the joint model, regardless of whether the factor loading random effects are in the joint model, happen by chance? We chose to investigate this question using a simulation study.

We simulated fire data using the Weibull baseline hazard function since this closed, parametric form is quite flexible. Therefore, $T_{ij} \sim \text{Weibull}\left(\lambda_j u_{ij} \exp(\boldsymbol{x}'_{ij}\boldsymbol{\beta}_j), \rho_j\right)$ where the hazard function has the same form as (4.2). For simplicity we assumed that there are no predictors in the models (i.e., $\boldsymbol{x}'_{ij} = \mathbf{0}' \implies \exp(0) = 1$) which yielded the lifetime distributions as $T_{ij} \sim \text{Weibull}(\lambda_j u_{ij}, \rho_j)$ where $\lambda_j > 0$ and $\rho_j > 0$. We chose to only look at the cases where the frailty terms have the one-parameter Gamma distributions as outlined in Table 4.1. The fire lifetime data was simulated using: (1) the FL form and (2) the FLI form. The results from (2) are provided here and a summary of the results from (1) are provided in Appendix B.

The one-parameter Gamma distributed simulated random effects are shown in Figure 4.15 where $\phi_u = \phi_v = 1/2$. The factor loading parameter γ was assigned to four different levels:

- $\gamma = 0.001$; extremely weak linkage,
- $\gamma = 0.1$; weak linkage,
- $\gamma = 1$; moderate linkage, and
- $\gamma = 10$; strong linkage.

Figure 4.16 illustrates the relationships between the simulated frailties of the detection phase and the action phase for the different cases. We see that the detection phase frailty dominates over the action phase frailty when there is a strong linkage coming from the factor loading parameter.

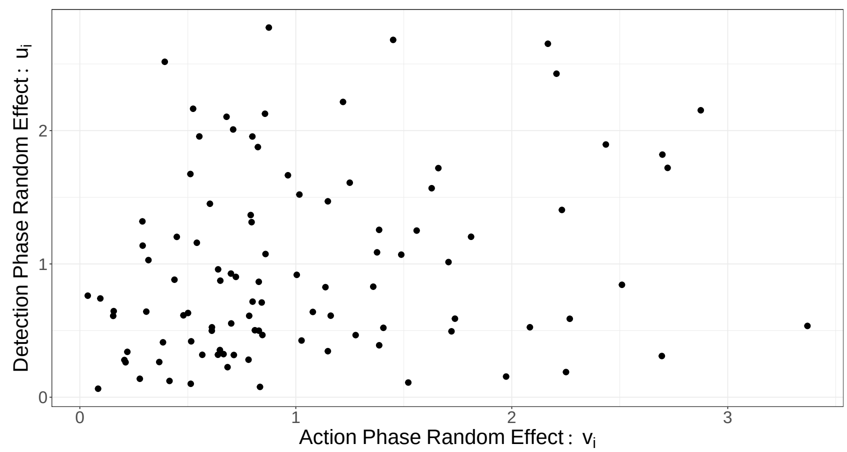


Figure 4.15: Scatterplot of the simulated random effects for the detection phase versus the action phase.

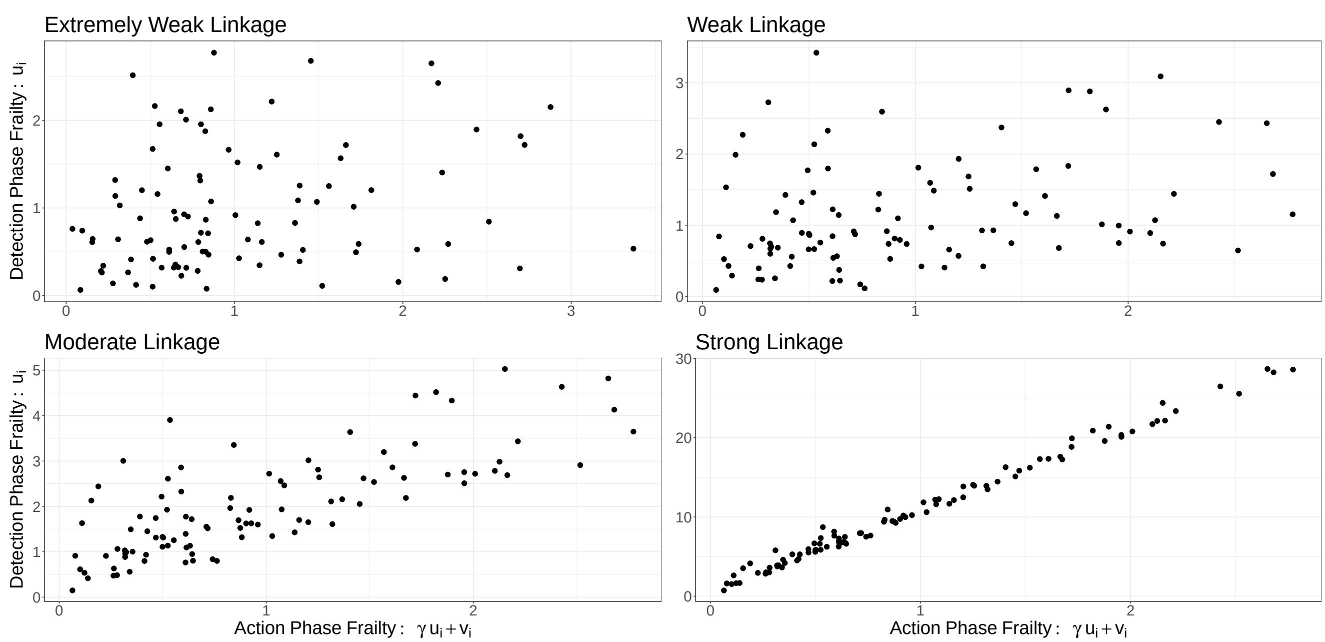


Figure 4.16: Scatterplots of the simulated frailties for the detection phase versus the action phase stratified by linkage.

We simulated a sample size of 100 detection and action lifetimes, using the different linkage cases, for one thousand runs or draws of the study (i.e., n=100 and draws=1,000). Within each draw and for each linkage case, we fit three joint models to the simulated data where the models either had an FL, FLI, or I form. For each draw, we calculate the ELPD estimates, the effective number of parameters, the information criterion estimates, and the ranking of the estimates, using the WAIC and PSIS-LOO methods. A model is ranked as 1 (i.e., the best) if it has the highest $\text{ELPD}_{\text{WAIC}}$ estimate or lowest WAIC estimate and ranked as 3 (i.e., the worst) if it has the lowest $\text{ELPD}_{\text{WAIC}}$ estimate or highest WAIC estimate. The same ranking order is used for the $\text{ELPD}_{\text{PSIS-LOO}}$ or LOOIC estimates.

A comparison of the ranked model forms, stratified by the ranking methods and the different linkages, is provided in Table 4.6. Our 'A, B, C' ranking order notation, where A is the best model form, B is the second best model form, and C is the worst model form, will continue throughout the rest of this chapter. For instance, the 'FLI, I, FL' ranking order means that FLI is the best model form, I is the second best model form, and FL is the worst model form.

The 'FLI, I, FL' ranking order of the model forms occurs the majority of the time across all linkage cases and for both ranking methods. The two ranking orders, 'FLI, I, FL' and 'I, FLI, FL', are the only ones that occur during the simulation, suggesting that the FL form is the least preferred model form which is most likely due to the fact that it is less flexible since it does not allow for the action phase to have its own random effect. The 'FLI, I, FL' ranking order of the model forms occurs more often for the WAIC ranking than for the PSIS-LOO ranking when looking at a specific linkage case. For instance, suppose we focus on the extremely weak linkage case. Here, we see that the 'FLI, I, FL' ranking order occurs 70% of the time when we rank by the WAIC estimates whereas it only occurs 54% of the time when we rank by the LOOIC estimates. However, the 'FLI, I, FL' ranking order of the model forms dominate as we move from the extremely weak linkage case to strong linkage case where it occurs 100% of the time.

Table 4.6: A comparison of the ranked model forms, stratified by the ranking methods and the different linkages.

	Ranked Model Forms			Draws	
	Best	2nd Best	Worst	Counts	Percentages
WAIC Rank					
Extremely Weak Linkage	FLI	I	FL	704	70%
	I	FLI	FL	296	30%
Weak Linkage	FLI	I	FL	694	69%
	I	FLI	FL	306	31%
Moderate Linkage	FLI	I	FL	887	89%
	I	FLI	FL	113	11%
Strong Linkage	FLI	I	FL	1,000	100%
PSIS-LOO Rank					
Extremely Weak Linkage	FLI	I	FL	538	54%
	I	FLI	FL	462	46%
Weak Linkage	FLI	I	FL	549	55%
	I	FLI	FL	451	45%
Moderate Linkage	FLI	I	FL	817	82%
	I	FLI	FL	183	18%
Strong Linkage	FLI	I	FL	1,000	100%

Some of the ELPD estimates are quite close to one another making the ranking of the "best" model somewhat contentious. Rather than only examining the WAIC and PSIS-LOO ranking orders, we also look into the differences between the estimates. We calculate the differences between the first ranked model and the second ranked model, along with the differences between the first ranked model and the third ranked model, for both the WAIC and LOOIC estimates. Figures 4.17 and 4.18 illustrate these differences across the different linkage cases. Additional density plots of the differences between ranked model forms are provided in Appendix B.

In general, we see that the differences between the FLI and I models are very small regardless of the ranking order or ranking method. More specifically, if the FLI model is ranked first using either ranking method, then the I model is closely behind it in second place. This also occurs if the I form is ranked first. But the differences between the first ranked model (i.e., FLI or I) and the third ranked FL model are often quite large. Interestingly, these differences increase as the factor loading parameter increases (i.e., as we move from extremely weak linkage to strong linkage).

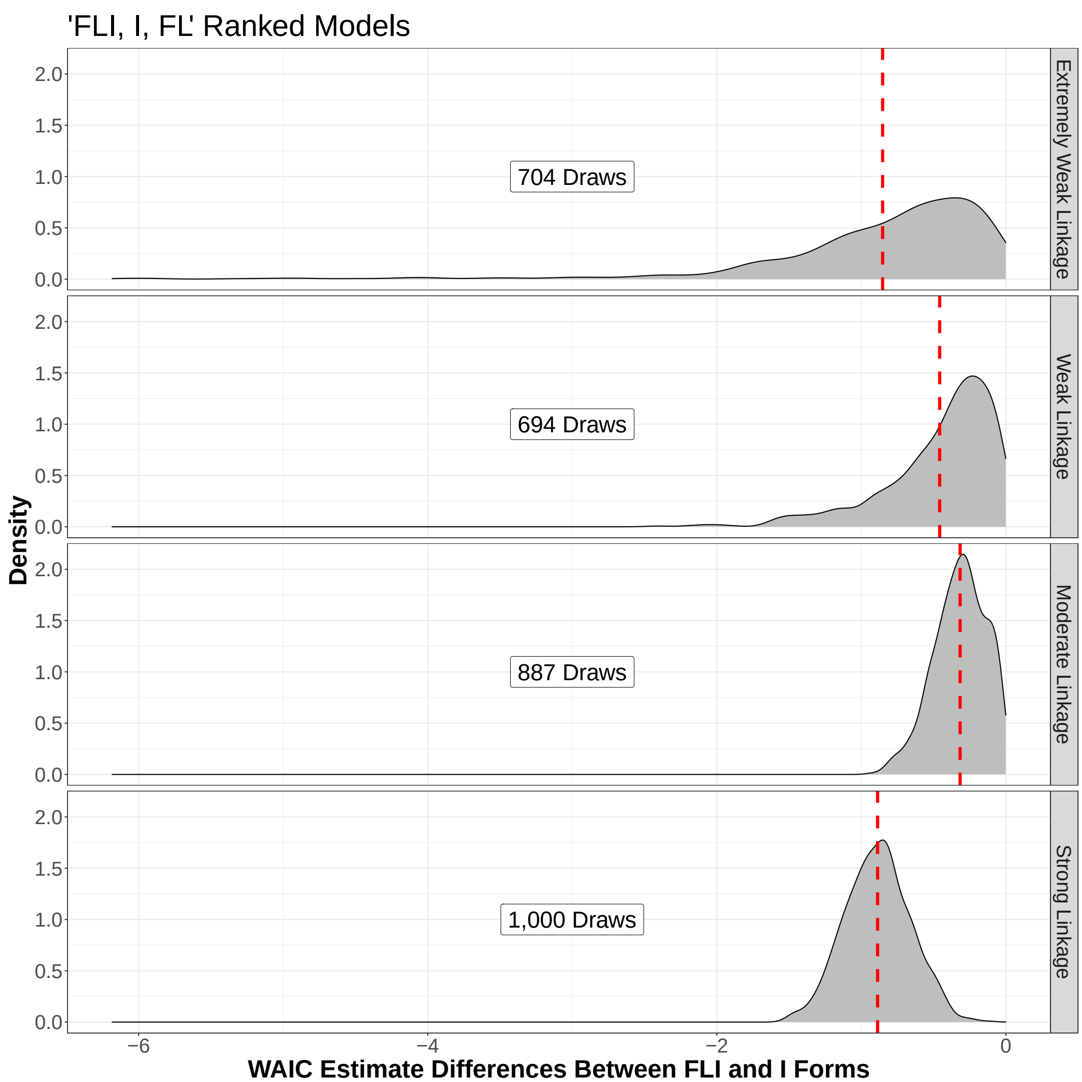


Figure 4.17: Density plots of the WAIC estimate differences between the FLI and I forms using the 'FLI, I, FL' ranked models. The red dashed line represents the respective means of the differences.

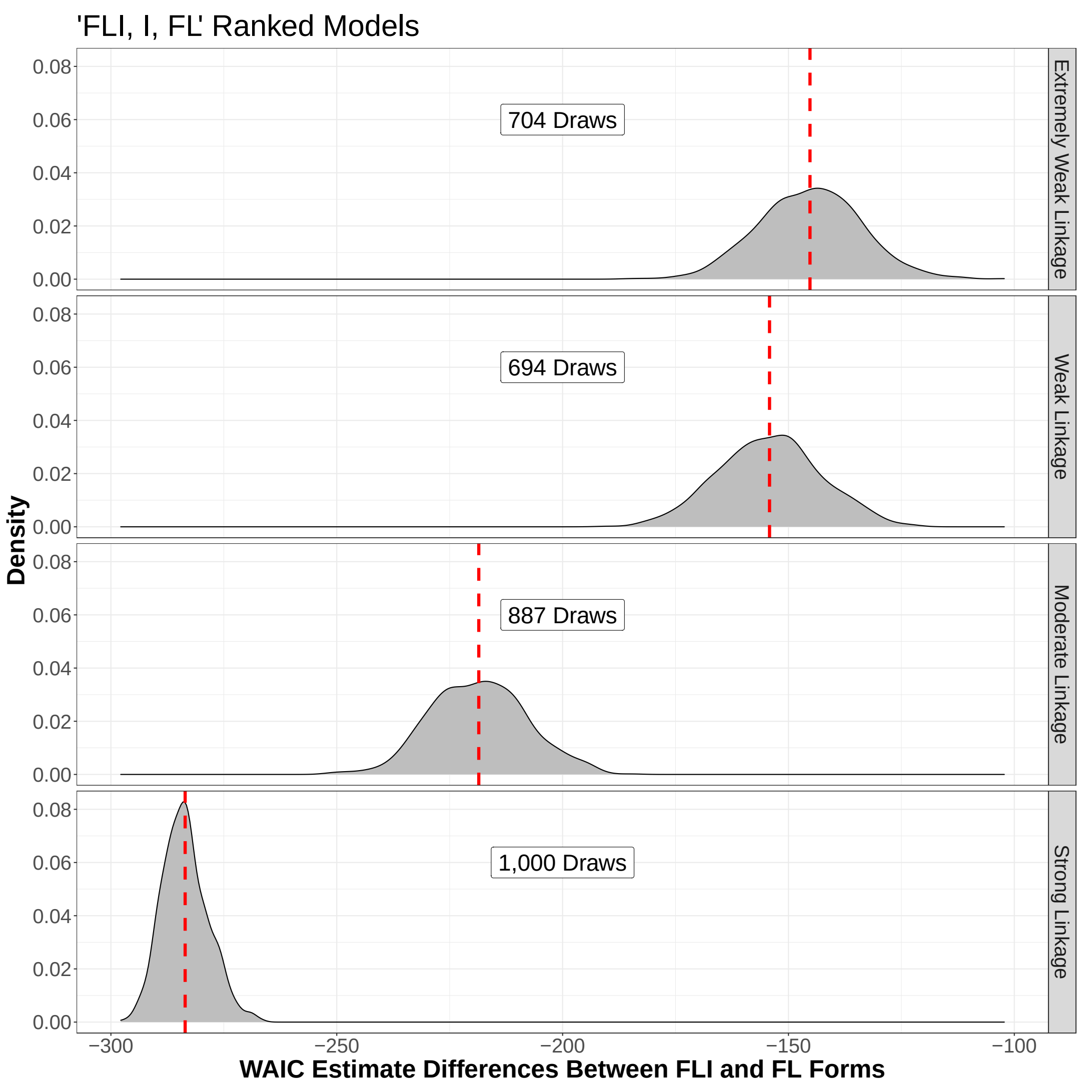


Figure 4.18: Density plots of the WAIC estimate differences between the FLI and FL forms using the 'FLI, I, FL' ranked models. The red dashed line represents the respective means of the differences.

4.4 Discussion

From our results we see that the preferred model for the human-caused fires has a Weibull baseline, FLI form, and Gamma frailties whereas the preferred model for the lightning-caused fires has a piecewise exponential baseline, FLI form, and Lognormal frailties. The analysis of data from the two fire causes agree on the form of the frailty term and their factor loading parameters, γ, are positive which implies that the detection phases are positively correlated with the action phases. In addition, note that the factor loading parameter is defined to be positive when using Gamma frailties as specified by the modelling framework when using that frailty distribution. The 90% credible interval for γ from the model for lightning-caused fires does overlap zero, implying that we cannot be certain of that relationship. The 'poorest' fitting model for both human- and lightning-caused fires is shown to be of S form using a Weibull baseline which suggests that a fire-specific random effect is necessary when jointly modelling two fire lifetime phases. The rest of our discussion will focus on the output from the two preferred models.

There are several factors that drive the fire lifetime phases from these models, particularly whether the report of and getaway to a fire occurred on the same day and whether the initial attack of a fire was successful. Most interestingly we see that the detection phase length is a driver of the action phase. Fires with longer detection phases are associated with longer action phases for both types of fires — thus it is crucial for the Ministry to prioritize the early detection of fires in Sioux Lookout to hopefully mitigate the length of time or amount of resources spent on actioning wildland fires.

Some interesting spatial and temporal trends appear when we investigate the posterior estimates for the detection and action frailties of both types of fires. Certain fires or areas, such as those along the highlighted corridor running along the highway between the Pickle Lake Attack Base and the Savant Lake Forward Attack Base for the human-caused fires in Figure 4.2, may require further investigation from the Ministry to determine why the action phases of those fires took longer after accounting for all other model effects.

Overall, it appears that the action phase lengths for both human- and lightning-caused fires may be getting longer over time. However, we must highlight that while these relationships are present in the data, there could be other important variables that drive these lifetime phases and could possibly change or influence the spatial and temporal trends shown.

Recall that both models use the FLI form which corresponds to having a shared frailty between the two phases and independent frailties for each phase. In fact, the FLI form and the I form tend to dominate the model ranking shown in Figures 4.1 and 4.5. Our simulation study investigates the ranking of the model forms further by comparing the WAIC and PSIS-LOO ranking methods for models fit using the FL, FLI, and I forms. Overall, we see that the two ranking methods are good at identifying when the FLI form is the true model form since it is most often: (1) ranked as the best model, or (2) ranked as the second best model. However, the two ranking methods are quite poor at identifying when the FL form is the true model form. We know that the fire lifetime phase durations are very heavily skewed to the right, especially for the lightning-caused fires, requiring more flexibility in the models of the lifetimes. Therefore, we argue that a flexible model which utilizes the FLI form — where the two phases of a fire lifetime are linked but still different from one another — is necessary when fitting models to fire lifetimes, and that this flexibility may be an important consideration when jointly modelling outcomes in other contexts.

There are some limitations to our analysis that we must note. First, the fire archive which stores the Sioux Lookout wildland fire data only provides the highest level of suppression used on a fire. This means that if a fire is monitored for a long time and then they choose to suppress it, it will be classified as a full suppression fire in the archive. Thus, there may be fires in our dataset that were originally monitored for a time until suppression action was taken which would increase the length of the action phase. Additional data would have to be collected to investigate the potential impact of this missing information.

Another limitation is the wildland fire data that was used for our modelling. We restricted our dataset to fully suppressed fires to ensure that the fires progressed through the various lifetime phases. However, monitored fires often have interesting lifetimes as well since they are allowed to burn under observation. Modelling both fully suppressed and monitored fire lifetimes is a future goal.

A third limitation of our modelling is the choice of baseline hazard functions. For the piecewise exponential baseline it would be helpful to perform a sensitivity analysis to determine the number of join points required, along with their locations, to ensure a reasonable number of fire lifetime phase durations within each interval. Also, employing a cubic B-spline basis expansion for the baseline hazard functions is another option that would provide flexibility in the models.

A major limitation of our modelling framework is the use of informative priors and hyperpriors for certain parameters. This may influence the results as presented in this chapter. These prior and hyperprior distributions were chosen at the time to achieve convergence across the 28 models fit in our analysis for the purspose of this thesis. Future work will explore the use of weakly informative prior and hyperprior distributions. Preliminary investigations into this have revealed that highly influential predictors, such as successful initial attack and same day dispatch, play a role in issues with convergence. Whether such variables should be used as predictors requires careful consideration in future modelling efforts using a joint-modelling framework for phases within a wildland fire's lifetime.

An important consideration of our modelling is that we used the detection phase as a predictor for the action phase and also linked the two phases together with the shared frailties (i.e., the factor loading form). Suppose that the detection and action phases are strongly linked to one another and behave in the same way (i.e., longer detection implying longer action). By fitting joint models where the detection phase is a predictor in the action phase, we are accounting for this linkage in a linear fashion and allowing the frailties to capture any non-linear effects between the two phases. However, if we remove the detection phase predictor then the factor loading parameter and frailties *must* pick up that linkage. In this case the interpretations of the effect that the detection phase has on the action phase becomes more difficult to explain to a layperson (e.g., fire management personnel) since a thorough understanding of frailties and non-linear effects are required. Although we plan to explore these ideas further with another simulation study using models with FL and FLI forms fitted with and without the detection predictor, we are cognizant of the change in interpretations and their potential complications.

Yet again, we must consider the stakeholders and end-users when developing our models and interpreting the results. The joint frailty models discussed in this chapter offer some interesting insights into the lifetime phases of fully suppressed wildland fires in Sioux Lookout, but it is crucial that this information is accessible for fire management personnel to understand. We took advantage of the principles of knowledge exchange to guide us in our process of collaboratively developing this research and the sharing of its results. For instance, Figures 4.19 and 4.20 plot the detection and action frailty densities by fire management strategy — strategies used by the Ministry that are vital to how wildland fires are managed in Ontario — which were created as part of this collaborative process. Without it we might not have realized that, under the appropriate response fire management strategy, we may be seeing longer action phases for both human- and lightning-caused fires. This may be a signal of "appropriate response" which may be of interest to the Ministry. Chapter 5 outlines how we integrated interdisciplinary knowledge exchange throughout our research project.

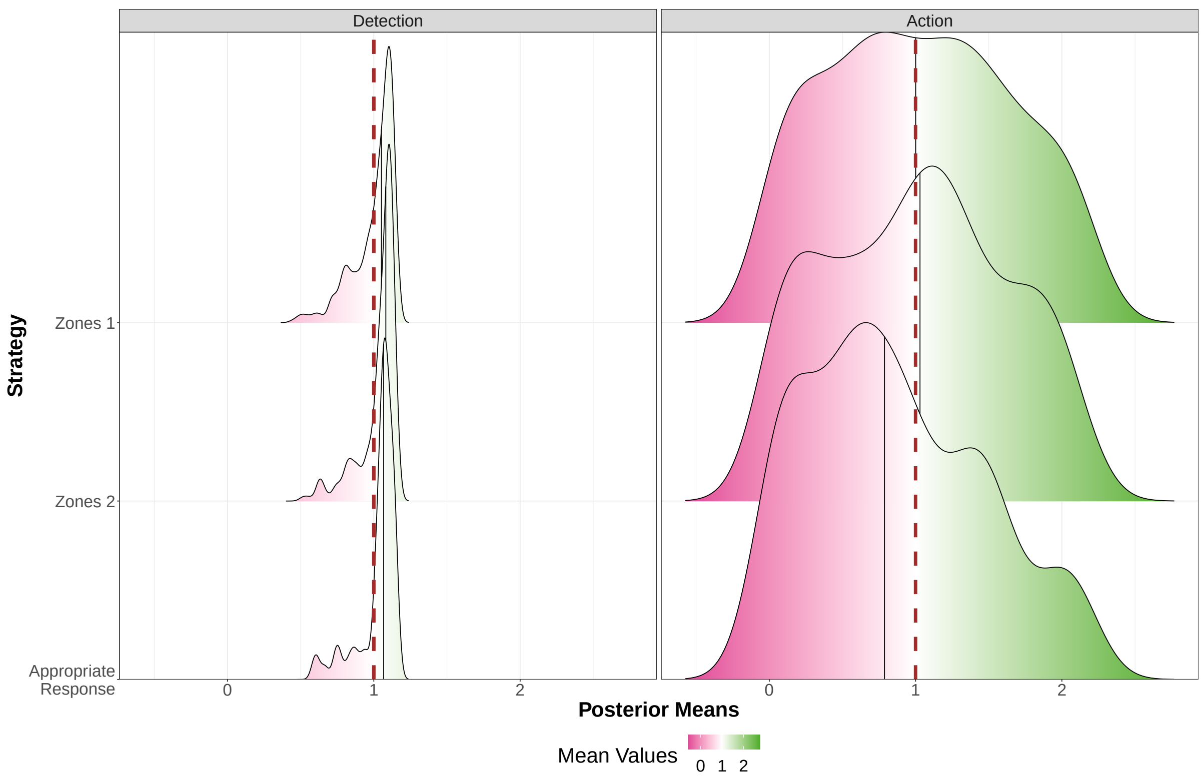


Figure 4.19: Plot of the human-caused posterior estimates for the detection and action frailties by fire management strategy.

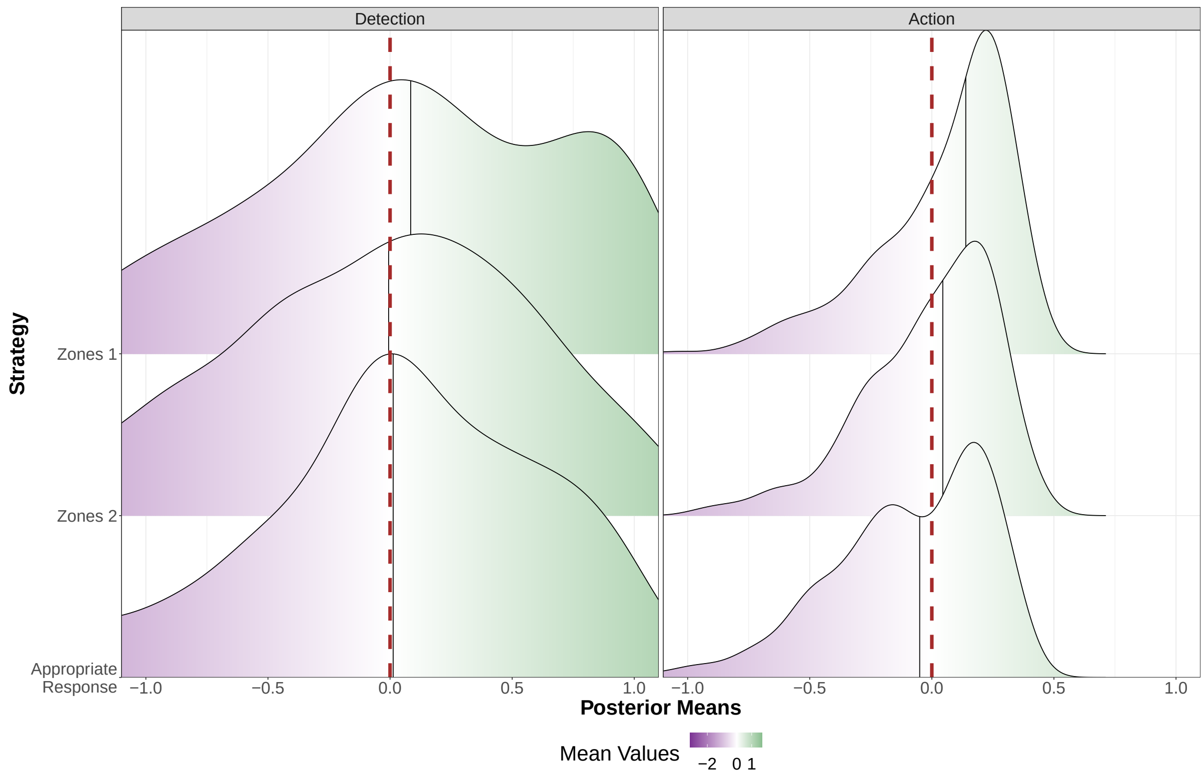


Figure 4.20: Plot of the lightning-caused posterior estimates for the detection and action frailties by fire management strategy.

Chapter 5

Reflections on Knowledge Exchange

5.1 Personal Reflections

In statistics, graduate students rarely get opportunities to practice true knowledge exchange as they work on their research. I have been fortunate that both my co-supervisors have encouraged, integrated, and prioritized components of knowledge exchange during my doctoral pursuits. My fire science research discussed in Chapters 3 and 4 benefited greatly from many interdisciplinary experiences that contributed to the knowledge exchange of this research and I describe below five key experiences (in bold) that played major roles in my development.

Over the past five years I have had several **tele/video-conferencing calls with fire management personnel** to assist with the creation and development of my research. This was especially helpful during the problem identification stage because it allowed us to zero in on a topic that was of interest to all parties. I attended a **workshop in February 2018 on "Wildland Fire Appropriate Response: Generating and Using Science" where the objective was to bring together a group of multidisciplinary researchers and fire management practitioners to**:

1. Discuss the barriers and opportunities to generate new data-driven science and

bring this into practical use for wildland fire management decision support;

2. Strengthen multidisciplinary teams collaborating on such problems; and
3. Provide opportunities for subject matter experts to design and work on solutions directly with fire management practitioners.

Participating in this workshop gave me greater insights into the concerns and needs of fire management personnel. Often, statistical training focuses on understanding the underlying theory of models and how to apply it when developing models that fit our data, but this workshop helped me to realize that a model is only as good as an end user's ability to understand it and its output. This experience was the first time that I ever considered how my work might be understood, used, or expanded upon by others and the necessity to incorporate these individuals into the research and development process.

For three weeks in July and August of 2019, I had the opportunity to **visit Dr. Meg Krawchuk's Landscape Fire and Conservation Science Research Group** at Oregon State University due to Western University's Science International Engagement Fund. My experience learning from and interacting with Dr. Krawchuk and her graduate students/colleagues was eye opening and thought provoking. I observed and performed ecological fieldwork on various burnt landscapes, clear cuts (i.e. areas with harvested trees), and arid grasslands in southwest and central Oregon which provided a better understanding of key ecological drivers and effects of wildland fires in the Pacific Northwest. This experience enabled me to expand my knowledge of wildland fire research, interact with international researchers from other fields, gain hands-on experience of ecological fieldwork, and increase my breadth of fire regimes outside of Canada.

I attended the **Wildland Fire Canada 2019 conference** held in Ottawa in November 2019 and presented a poster on the initial findings of our wildland fire lifetime research (see Appendix C). It was very helpful to participate in this conference because of the discussions I had with fire science researchers and fire management personnel about these early findings and ways that I might strengthen my research methodologies.

I was also fortunate to discuss my research findings and methodologies with my peers across various disciplines through the completion of a **Collaborative Specialization in Environment and Sustainability** from September 2018 to April 2021, offered by the University of Western Ontario's Centre for Environment and Sustainability. The collaborative specialization served as an interdisciplinary enrichment program for graduate students whose research coincided with the environment and sustainability. During our weekly seminars we had the opportunity to: (1) share our research findings and offer suggestions through an interdisciplinary lens; (2) create solutions to community projects that addressed specific environmental problems; and (3) develop an annual collaborative conference on environment and sustainability called EnviroCon.

As our research progressed to the stage where we were interpreting results and gleaning insights from our findings, we reached out to Colin McFayden from the Ministry to continue our knowledge exchange journey. As a collaborative team, we are in the process of developing a communiqué for Ministry members, specifically those from the Sioux Lookout District. The process of developing this brief report of our wildland fire research, for an audience with backgrounds in fire science and management rather than statistics, has been extremely helpful in understanding our own work. For example, we were required to explain the concepts of “frailties”, “joint models”, and “longer/shorter phase lengths *after accounting for all other model effects*” in a few short pages without using technical terms. This deliverable is analogous to the popular 3 Minute Thesis[1] whereby having to explain something to someone else, especially in a shortened format, helps you understand it better yourself - which was certainly the case for us. We also had several interesting discussions about the colour schemes used in our visualizations; green, yellow, orange, and red already have specific associations due to the Canadian

[1]The 3 Minute Thesis (3MT) is a university-wide academic competition for Masters and Doctoral students in which participants present their research and its wider impact in 3 minutes or less to a panel of non-specialist judges.

Forest Fire Danger Rating System (Natural Resources Canada, 2021b).

Overall, the incorporation of these experiences focused on the knowledge exchange of my research played a crucial role throughout my doctoral journey. They helped me to become a well-rounded, collaborative researcher who prioritizes the needs, suggestions, and ideas of stakeholders, end-users, and any other members involved in a project. The breadth of my professional training and development as a statistician was greatly expanded from these opportunities and I am truly grateful for them.

Knowledge exchange is crucial for any practicing statistical scientist, yet statistical training often does not focus on this. As a graduate student, I prioritized seeking out training programs and experiences to develop and expand my teaching-related skills. For instance, I was given the opportunity to teach at the university-level early on in my doctoral pursuits which only furthered my interest in becoming an educator. My passion for teaching also led me to consider the idea of knowledge exchange outside of the wildland fire science context, specifically into the realm of education for the mathematical and statistical sciences. Throughout my graduate studies I engaged in professional development related to education by participating in various workshops, seminars, and programs offered through Western's Centre for Teaching and Learning. Eventually, this led to formal Scholarship on Teaching and Learning, namely developing a teaching-related training and development program — focused on active learning (i.e., an embodiment of knowledge exchange) — for graduate teaching assistants in Western's School of Mathematical and Statistical Sciences. Chapter 6 provides details about this program and the associated study that we performed.

Chapter 6

Investigating Graduate Teaching Assistant Training and Development in Western University's School of Mathematical and Statistical Sciences

Education is one of the main pillars in the discipline of statistics, as evidenced by the attention paid to statistical education by both the American Statistical Association (ASA) and the Statistical Society of Canada (SSC). Here, we focus on the teaching-related training and development of graduate students within the statistical and mathematical sciences, specifically spotlighting the training of graduate teaching assistants on active learning techniques. Such techniques allow for knowledge exchange — an interactive and collaborative flow of information — between students and instructors (or graduate teaching assistants) and between students and students, within different learning environments (e.g., the classroom).

6.1 Literature Review

Should we ask or should we tell in the Science, Technology, Engineering, and Mathematics (STEM) classroom? Freeman et al. (2014) sought to answer this question by performing a quantitative data analysis to determine how constructivist (or learner-centred) methods like active learning versus exposition-centred (or instructor-centred) methods like lecturing impacts student performance in undergraduate STEM courses. They meta-analyzed 225 studies in both published and unpublished literature consisting of 158 independent comparisons with data on student examination performance and 67 independent comparisons with data on failure rates. Their results showed that incorporating active learning in such courses increases student examination performance and that not doing so (i.e., exclusively lecturing) increases student failure rates by 55%. These results suggest that the student-centred approach to teaching (O'Neill and McMahon, 2005) may lead to increases in student performance which strengthens the call to include more active learning in undergraduate STEM courses.

It is important to note that an instructor-centred (or teacher-centred) mindset consists of: knowledge being transmitted only from instructive, passive student participation, the lecturer being the leader and authority in the classroom, assigning few assessments meant solely for grading, and one-dimensional assessment methods where the emphasis is placed on learning correct answers. In contrast, a learner-centred (or student-centred) mindset consists of: knowledge being constructed by students, active student participation, the lecturer being the facilitator or partner in the classroom, assigning many assessments meant for ongoing feedback, and multi-dimensional assessment methods where the emphasis is placed on developing a deeper understanding. The academic culture using the instructor-centred approach is often competitive and individualistic, whereas it is collaborative and supportive for the learner-centred approach. For more details on these topics see O'Neill and McMahon (2005) and Wright (2011).

There is a need to expand the pedagogical training of graduate teaching assistants

(GTAs) using teaching development programs. Reeves et al. (2016) noted that GTA teaching professional development programs directly impact GTA cognition which then impacts GTA teaching practices and, by association, undergraduate student outcomes. Such programs, along with their designs and benefits, are explored below.

First, Gilmore et al. (2014) examined four variables that impact GTAs' teaching orientations: mentorship, training for teaching, teaching experience or teacher development, and research experience. They hypothesized that each factor is positively associated with GTAs' teaching development programs that become more student-centred over time. Most interestingly, they found that neither the duration of prior teaching experience nor the duration of research experience was significantly related to a change in teaching orientation over time and suggested that the quality of those experiences was more important than the length.

Campus-wide graduate teaching development programs at the University of Western Ontario and the University of Windsor were assessed by Dimitrov et al. (2013) who compared the impact of such programs of varying duration and examined how GTAs apply what they learn in short (i.e., one day) and long (i.e., 20-40 hours) programs when they teach in their disciplines. They employed a mixed-method study design[1] that involved self-reported measures of participants' attitudes to teaching and teaching self-efficacy before and after the programs, along with focus group interviews occurring four months after program completion. The qualitative data from the interviews provided a more detailed description of how GTAs use what they learn from teaching development programs.

Possible changes in participants' instructor-centred and student-centred attitudes to teaching in the study by Dimitrov et al. (2013) was measured using the revised Approaches to Teaching Inventory 22-item standardized measure (Trigwell et al., 2005). This measure is comprised of two scales: the Information Transfer/Teacher-Focused scale

[1] A mixed-method study design utilizes both qualitative and quantitative data collection and analyses.

(e.g., "In this subject, students should focus their study on what I provide them.") and the Conceptual Change/Student-Focused scale (e.g., "I set aside some teaching time so that the students can discuss, among themselves, key concepts and ideas in this subject."). A high mean score on that scale reflects the respondent's focus on changing students' ways of thinking about a subject and recognizes the active role that students play in constructing their own knowledge. Teaching self-efficacy (i.e., the confidence GTAs feel in executing various teaching behaviours or duties) was also measured using the Teaching Assistant Self-Efficacy scale (Boman, 2008). Participants rated their confidence in performing various GTA duties on a five-point Likert scale (from 1 = not confident to 5 = completely confident).

The participants in the study by Dimitrov et al. (2013) highlighted that the most useful sessions in the programs were ones where they gathered concrete teaching strategies to use later and ones where they heard about personal experiences of other GTAs and faculty. The results of their samples at the University of Western Ontario showed that participants' student-focused approach to teaching increased throughout both short- and long-term programs. However, the participants of the shorter programs were more student-focused than the participants of the longer programs. Dimitrov et al. (2013) suggested that these participants of the shorter program are more aware of student needs since they are typically younger, have recently completed their undergraduate degree, and have less teaching experience.

Rivera (2018) sought to answer the research question: "What does a Summer Institute (SI) on STEM GTA pedagogy and the experience of first year GTAs reveal about STEM GTA perceptions concerning their roles as a GTA?" Their 5-week intensive summer training session for STEM GTAs was created by an interdisciplinary group of faculty consisting of classes like: STEM Methods, STEM Literacy, Teaching Labs, and Professional Communications. Rivera argued that, although training GTAs in pedagogy increases the likelihood they will try various pedagogical methods to improve their teaching, there is a gap regarding what GTAs already know and believe about teaching. Filling in that gap will help discern how to better prepare GTAs for their roles and, by extension, improve STEM undergraduate education. Rivera concluded that the SI improved the teaching confidence of all GTAs (especially international students) and that it helped shift the instructor-centred mindset to a learner-centred one.

A special section of *The American Statistician* (2005, Vol. 59, Issue 1) highlighted strategies for preparing GTAs for teaching across different statistics departments in the United States. Moore (2005) expressed that "the issue before us [the statistics community] is how to help graduate students learn the craft of teaching" (p. 1). These strategies included courses in teaching statistics (Gelman, 2005; Harkness and Rosenberger, 2005), weekly meetings (Harkness and Rosenberger, 2005), mentoring (Froelich et al., 2005), and progressive training providing immersion in a departmental culture (Birch and Morgan, 2005), all of which encourage student-centred teaching.

Several courses in teaching statistics have been developed at various institutions. For example, Gelman (2005) developed "The Teaching of Statistics at the University Level" course at Columbia University, which was required for all first-year Ph.D. Statistics students and was offered in the fall (i.e., first) term of their degree. Its objective was to boost their graduate students' confidence and effectiveness in teaching statistics with active participation and to ensure that they would be prepared to handle the practical difficulties that arise in teaching. Garfield and Everson (2009) created a graduate-level course called "Becoming a Teacher in Statistics" at the University of Minnesota, originally using a face-to-face setting then converting it to an online course. The course was designed to help students understand and align their teaching with the ASA-endorsed GAISE recommendations.

Green (2010) performed a study using focus groups, e-mails, and an interview to collect qualitative data on ten statistics teaching assistants to determine the experiences they had at the University of Nebraska-Lincoln while teaching the STAT 218: An Introduction to Statistics course. Their results determined that novice statistics GTAs need specific direction to help them develop their pedagogical content knowledge for teaching statistics and that statistics departments need to prioritize GTA preparation and training.

Justice et al. (2017) developed an online survey, called the *Graduate Student Statistics Teaching Inventory (GSSTI)*, to better understand the preparation, teaching beliefs, and teaching practices of statistics GTAs in the United States. Responses from 213 GTAs enrolled at 38 Ph.D.-granting institutions for all major regions across the United States were collected. The results showed that many GTAs had not yet learned about student-centred teaching principles like active learning activities since only 40-55% of respondents had indicated that they had learned about these topics. It also highlighted that the GTAs had little consensus about their pedagogical beliefs since none of the survey question items related to the delivery of course content (e.g., lecture, activities, and small group work) reached 60% agreement and over 10% responded as ‘undecided’ to each of the questions.

Justice (2020) summarized six empirical studies conducted in the United States for preparing or training GTAs in teaching statistics. They found that “there appears to be consensus that many GTAs in statistics departments need more knowledge, preparation, and support as they fulfill their teaching roles” (Justice, 2020, p. 336). Justice (2020) also reviewed training of GTAs in other disciplines and discovered that the following components of GTA training programs are helpful in training GTAs: teaching observations, mentoring, and participation in a community of practice. A culmination of ten recommendations for GTA professional development programs for teaching statistics is offered by Justice (2020). These recommendations include providing GTAs with opportunities to develop pedagogical knowledge for teaching statistics and establishing formal interactions between experts and novices (e.g., faculty and GTAs, or senior and novice GTAs).

Many issues arise when GTAs are not properly trained for their roles; Gardner and

Jones (2011) discussed some of these. They commented on previous studies which showed that GTAs felt student success had less to do with the course and more to do with a students' ability or motivation; a clear misconception about how undergraduates learn and therefore how they should be taught. This mindset perpetuates the idea that teaching is centred on transmitting knowledge from the instructor to the student and is, therefore, less concerned with developing learner autonomy and independence. Their research also addressed issues with motivating GTAs to attend teaching development training sessions. Most science departments, and even some institutions, place little emphasis on pedagogical training and offer GTAs few rewards for improved teaching which undervalues the need to attend teaching development training and makes it acceptable to be neglected altogether. Gardner and Jones (2011) argued that "teaching beliefs and values are reflected in the behaviours of the institution, department, and the faculty and these ideals are [then] reproduced in graduate students" (p. 38).

Crowe (2019) focused on determining what factors affect STEM GTAs' perceptions of pedagogical training and whether these factors influence their buy-in to such training. They noted that previous factors identified in the literature included: departmental demands, the perception of research as more valuable than teaching, previous teaching experiences, self-efficacy, attitudes about teaching in general, and career goals. Crowe discovered two additional factors that may influence GTA buy-in to pedagogical training: the effect of pedagogical training on overall GTA learning and a social commitment to students. These additional factors may prove useful for encouraging buy-in from STEM GTAs. For example, if post-secondary institutions, faculties and departments present teaching as a way of enacting social responsibility then it might inspire GTAs to participate in training opportunities.

6.1.1 The Need for Research on GTA Training

GTAs play critical roles in educating the next generation of professionals. They influence both the undergraduate students they teach and the teaching effectiveness of future faculty. Gardner and Jones (2011) argued that the best point to implementing reform in the quality of undergraduate education is with GTAs since: (1) their increasing roles and responsibilities provide an opportunity to further reach teaching objectives set by universities, and (2) their hands-on training can be scaffolded with pedagogical training prior to faculty appointments to break the cycle of mentors with little or no formal training of this kind. GTAs must adopt effective pedagogical practices early in their careers since early teaching experiences tend to establish enduring teaching skills and approaches (Gilmore et al., 2014).

Training helps GTAs reflect on the expectations they place on their undergraduate students and the expectations that students place on them, reinforcing an environment of care, respect, and empathy for one another. It can also help to reduce the teaching anxiety felt by already over-burdened and stressed GTAs (Williams, 1991). The dedication of departmental or faculty resources to teaching training programs provides tangible proof that the department or faculty values this crucial aspect of graduate programs (Pentecost et al., 2012). However, until teaching, and the training of effective teaching, are prioritized as beneficial skills and valuable commitments of time, GTAs are likely to find it challenging to maintain this development on their own.

Such literature outlines an important gap that we aim to fill: increasing the teaching roles and responsibilities of GTAs requires increasing the research into the training they receive (if any). This gap was noted by Gardner and Jones (2011) who stated:

> "Although there are programs that exist to prepare science GTAs to be more effective instructors, there is a dearth of primary research on the subject. What do programs geared toward developing the teaching skills of science GTAs look like? More importantly, how effective are these programs at chang-

ing misconceptions and beliefs about science teaching and learning of science GTAs?" (p. 34)

Our main goal is to contribute to these conversations by performing a study that seeks to answer the following research question: **How does participation in a discipline-specific teaching development program on active learning for Graduate Teaching Assistants (GTA) in mathematics and statistics, offered by their School of Mathematical and Statistical Sciences, impact their perceptions of teaching?**

6.2 Methods

6.2.1 Workshop Development

A 1.75-hour long workshop on active learning, entitled "Active Learning in Math & Stats: Benefits, Limitations, and Practical Strategies for Implementing Active Learning Activities in Undergraduate University Mathematics and Statistics Courses", was developed for the purposes of our study. The initial version of this workshop was created in 2018 as part of a capstone project for the Advanced Teaching Program offered by the Centre for Teaching and Learning at the University of Western Ontario. A final version of the workshop was created in March 2021 in collaboration with Lisa Aikman (Education Developer of GTA Programs at Western's Centre for Teaching and Learning).

The goal of the workshop was for participants — namely, GTAs in mathematics and statistics — to gain a better understanding of active learning teaching methods (see Chapter 2.10 for more details). The learning outcomes for the workshop appear in Figure 6.1. By developing tangible examples of active learning activities applied to participants' home disciplines — in collaboration with their peers — and discussing their benefits and limitations, we envisioned that these activities would help GTAs to start thinking about different ways to implement these ideas in their labs, tutorials, office hours, and courses.

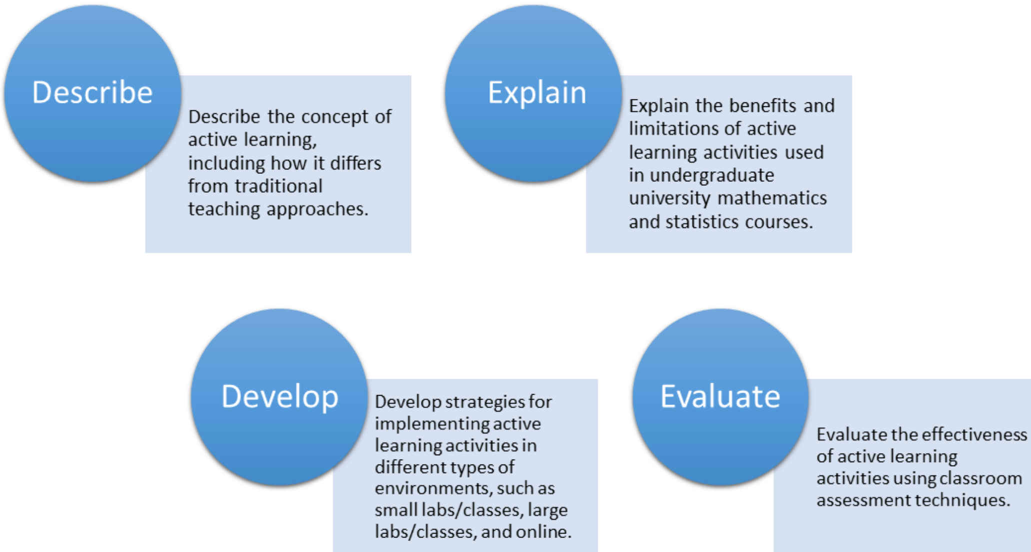


Figure 6.1: Learning outcomes for the workshop.

6.2.2 Study Design Outline

Study participants were asked to complete three tasks: a pre-workshop questionnaire, a workshop on active learning, and a post-workshop questionnaire. We use this pre-post self-report survey framework to investigate our research question.

The pre-workshop questionnaire was administered as an online survey through Qualtrics (`https://mysurveys.uwo.ca/`) — a web-based survey tool to conduct survey research — taking approximately 15 minutes to complete. The survey asked participants for demographic information, along with written responses and multiple-choice questions related to their GTA experiences and experiences with active learning. See Appendix D for the pre-workshop questionnaire.

The workshop was offered to all GTAs in research-focused graduate programs (M.Sc., Ph.D.) in the departments of Mathematics and Statistical and Actuarial Sciences on September 16, September 23, and October 7, 2021. The workshops were held in an active learning space[2] similar to the Western Active Learning Spaces (`https://www.uwo.ca/wals/`). We offered three sessions of the same workshop on the different dates for logistical reasons such as scheduling conflicts and room capacity constraints.

The post-workshop questionnaire was also administered as an online survey through Qualtrics taking approximately 10 minutes to complete and was sent to workshop participants one day after attending their workshop session. Appendix D contains the post-workshop questionnaire, which includes the same written responses and multiple-choice questions as the pre-workshop questionnaire, with two additional questions focusing on how the participant felt after completing the workshop.

It is important to note that pre- and post-workshop questionnaires were not standardized. The questions were chosen by examining the surveys from previous GTA studies and "approaches to teaching" studies, as found in the literature. Several multiple-choice questions are Likert scale questions (Johns, 2010) that utilize five-point ratings. Significant time was spent crafting the wording of the questionnaires, asking several colleagues with graduate training in the statistical sciences to review them and provide feedback on clarity to ensure that the questions were not biased, leading, or unclear, and revising them.

The target population of study participants were current and future graduate students in full-time research-focused programs (M.Sc., Ph.D.) within the School of Mathematical and Statistical Sciences at the University of Western Ontario. Participants did not need to have any previous experience working as a GTA. We offered an incentive where participants chose to be entered into a draw to win one of two $250 gift cards to Best Buy; participants were entitled to one ballot for each survey they participated in, for a total of two ballot entries in the draw.

Our study received approval from the University of Western Ontario's Non-Medical

[2] An active learning space (or classroom) is designed to support teaching and learning in an atmosphere conducive to engaging students actively in their own learning. They are equipped with tools that promote active and collaborative learning.

Research Ethics Board (NMREB) on July 20, 2021. Email recruitment began two weeks prior to the start of the first workshop session and ended one week before the last workshop session, using the mailing lists for each department's graduate students. A Program Coordinator assisted with emailing all the potential participants. We also reached out to the Graduate Chairs and faculty members in the respective departments to ask them to encourage their graduate students to attend the workshop.

6.3 Analysis and Results

6.3.1 Review of Workshop Sessions

An outline of the workshop including the associated summary, annotated bibliography, schedule, presentation strategies, and more information are provided in Appendix D. A total of five active learning activities were utilized throughout the workshop (e.g., Dotmocracy, Think-Pair-Share, etc.) as examples of potential activities that GTAs could use. Please review the workshop outline in Appendix D for descriptions of the activities. The workshop sessions ran rather smoothly with most participants picking up new ideas that they were eager to implement during their office hours, tutorials, and labs. In total there were 12 workshop participants: 4 attended the first session, 5 attended the second session, and 3 attended the third session.

The motivating activity used in the workshop was a Dotmocracy where participants were asked to vote by placing dot stickers under the statements they agreed with. Some participants did not know what active learning was at the time, so they could not agree with any of the statements that we used in the workshop. For example, one participant remarked in the first session that they did not know what active learning was, but were not skeptical about it either. Our suggestion for the next iteration of this workshop would be to use a more diverse range of statements for this activity so that everyone can vote in some way. A simple fix may include statements like "I do not know what active learning is" or "I have never tried active learning before". Most of the statements did receive dot sticker responses throughout the three sessions. Interestingly, the statement "I've taken an undergraduate math/stats course that included active learning" was not acknowledged as true by any of the participants until the last workshop session. This might be because participants did not know what active learning was at the time, even though they may have already encountered it in their math/stats courses, or because they genuinely had not encountered it in these courses.

A Quescussion[3] activity was used in the workshop as an example of an active learning activity that GTAs can use. In the first session, the participants struggled with creating questions and needed assistance from the workshop facilitator. During the second session, no participants asked any questions but instead sat in silence for a few minutes. Rather than moving on to the next activity, the facilitator turned the situation into a discussion of what to do when an activity does not work (i.e., what to do when students do not participate in an activity), which resulted in a helpful "meta-moment". During the third session this activity went extremely well since all participants asked very insightful and interesting questions and they actually reached the goal of the activity. The success of this activity in the last session may be due to how the facilitator explained the instructions (i.e., they became better at explaining the activity instructions with each session) or it may have been due to the knowledge and experience of those participants in the room.

6.3.2 Review of Questionnaires

Identifiable information collected from study respondents included citizenship and residency status (e.g., Canadian/Permanent Resident or International student), degree, program, year of study, whether English is their first language, age range, gender, and a unique personal identifier used to link the pre- and post-workshop questionnaires. We

[3] A Quescussion is a discussion conducted entirely by asking questions. One person (often the instructor or GTA) will ask a challenging question and a discussion of that question occurs where students can only respond by asking further questions.

collected 11 total responses with corresponding pre-post data, resulting in a high response rate ($\approx 92\%$) but a low workshop participation rate ($\approx 12\%$). Due to the low number of respondents, the rest of this section focuses on an exploratory data analysis of the collected responses.

Figures 6.2-6.5 show the demographics of our 11 study respondents. We see that 82% (9/11) of the respondents are enrolled in a Ph.D. program with years of study ranging from year one to year 5+. Respondents self-identified their gender with 36% (4/11) identifying as female and 64% (7/11) identifying as male, with ages ranging from 20 to 34 years old. Most of our respondents (82%; 9/11) are international students whose first languages are not English, whereas all the domestic/permanent resident students have English as their first language.

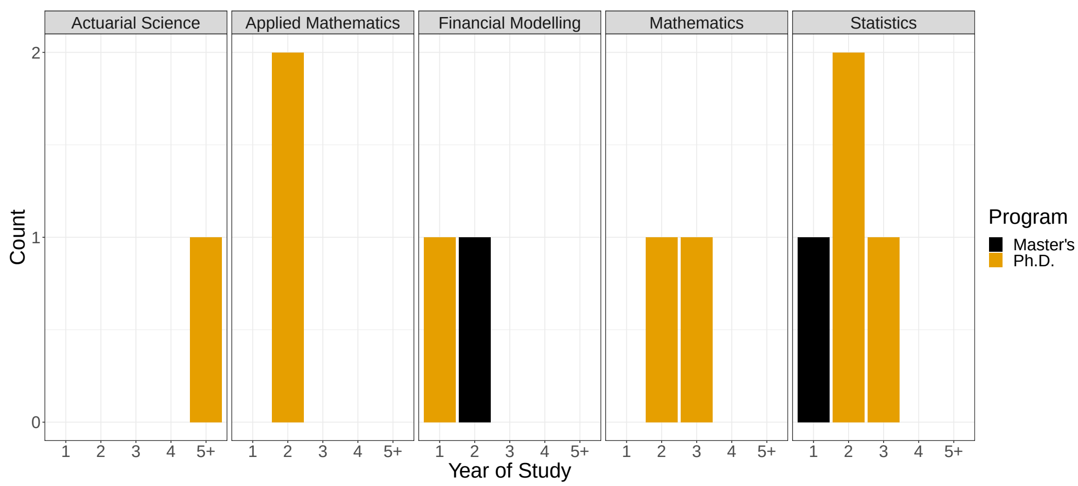


Figure 6.2: Bar chart of study respondents stratified by program, degree, and year of study.

Interestingly, in Figure 6.4 we see that three respondents have no previous GTA experience, whereas the rest of the respondents have at least 3 academic terms of experience (an academic term is considered four months). The three individuals who have no previous GTA experience includes the two Master's students and one Ph.D. student who is in their second year of study. It illustrates that Master's students and, potentially, early-career (i.e., in their first two years of study) Ph.D. students may not have previous

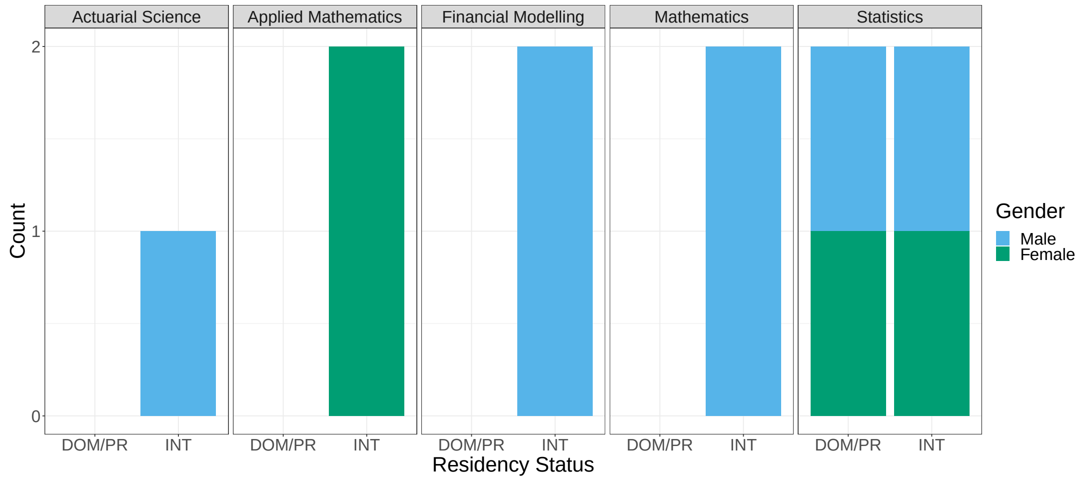


Figure 6.3: Bar chart of study respondents stratified by gender, degree, and residency status (DOM/PR represents domestic/permanent resident; INT represents international).

GTA experience when they are assigned to take on the role of a GTA.

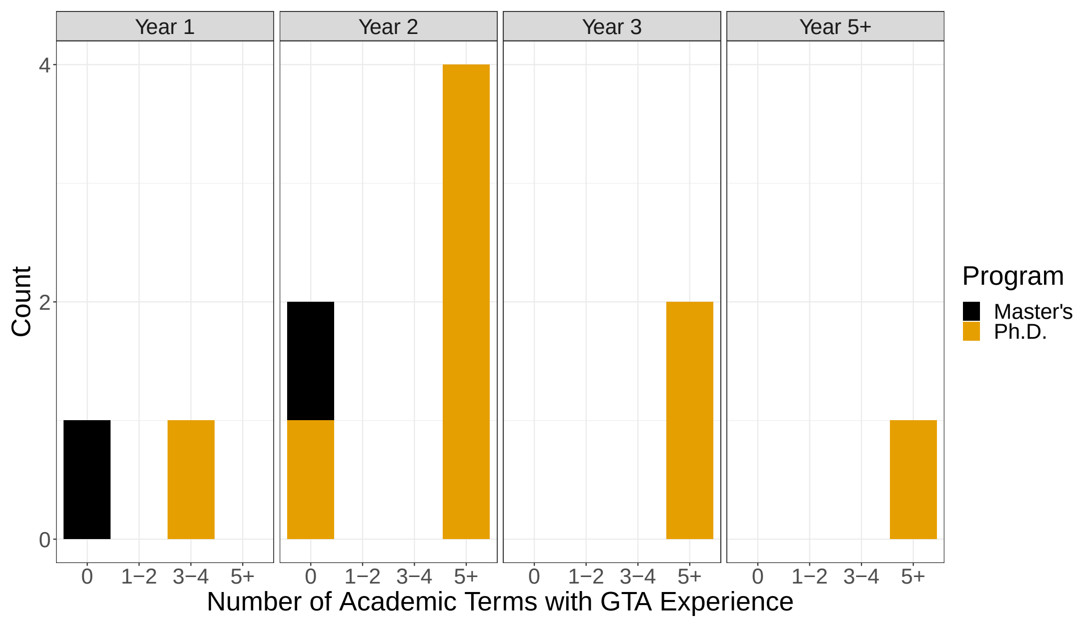


Figure 6.4: Bar chart of study respondents stratified by their program, year of study, and previous GTA experience.

Figure 6.5 shows whether the respondents have had any experience with different types of professional development activities related to teaching. We see that the majority (73%; 8/11) of respondents have not attended a conference, roughly half (6/11) have attended a short workshop of 1-5 hours (e.g., Future Prof Series), 64% (7/11) have attended a medium workshop of 1-2 days (e.g., Teaching Assistant Day), and only 36% (4/11) have attended a longer workshop of 3-10 days (e.g., Teaching Assistant Training Program or Advanced Teaching Program). Only two respondents attended a summer or semester-long course (e.g., SGPS 9500 course) and none of the respondents had any other teaching-related professional development.

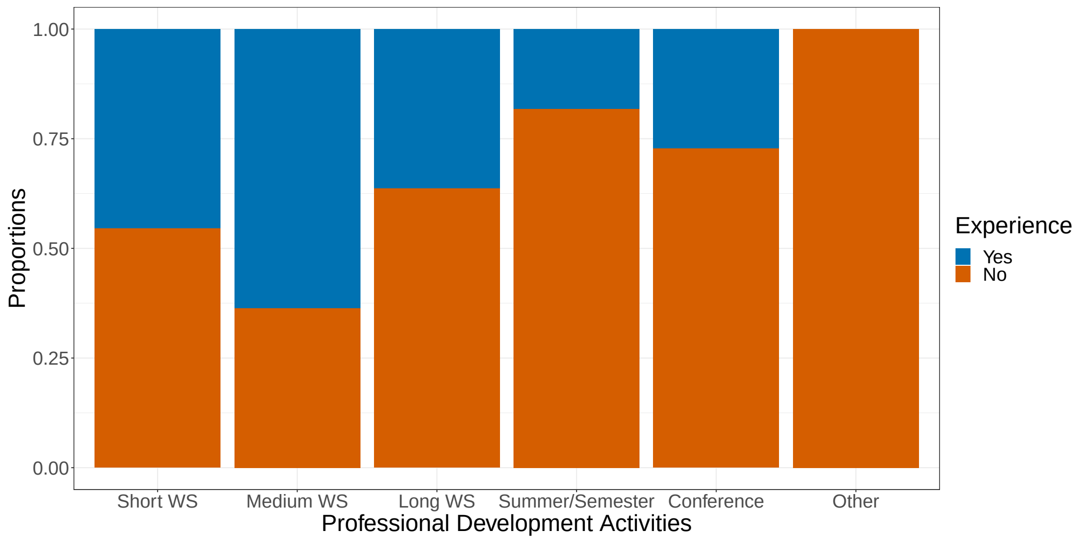


Figure 6.5: Stacked bar chart of the proportions of student respondents' previous professional development activities that are related to teaching.

The following figures (Figures 6.6-6.14) correspond to the workshop-related questions from the questionnaires, highlighting any possible changes in a respondents' response from the pre-workshop questionnaire to the post-workshop questionnaire. In Figure 6.6 we see that two respondents were slightly encouraged to consider using active learning strategies, one was slightly discouraged, and eight remained unchanged in their views. For example, Respondent 6 went from never considering these strategies to sometimes considering them.

Figure 6.7 corresponds to respondents' familiarity with active learning strategies used in their discipline (that is mathematics, statistics, actuarial sciences, or financial modelling). We see that four respondents were slightly more familiar with active learning strategies, one was slightly less familiar, and two remained unchanged in their familiar-

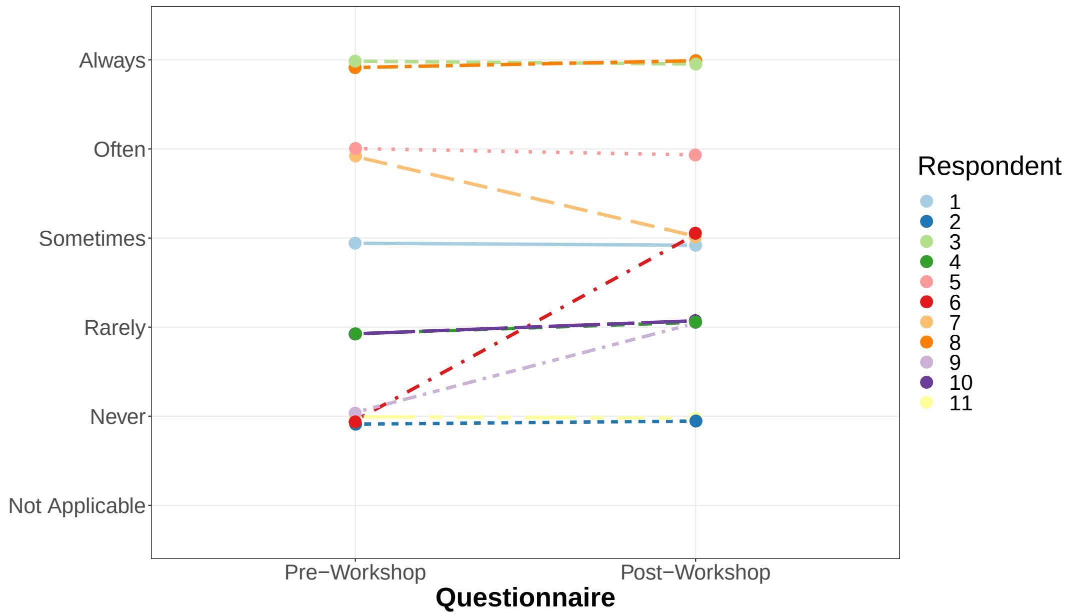


Figure 6.6: Connected scatterplot of responses to the question "Fill in the blank: As a GTA, I have (blank) considered using active learning strategies in labs, tutorials, and/or office hours."

ity. Interestingly, Respondent 1 went from being moderately familiar with active learning strategies before the workshop to slightly familiar with them after attending the workshop. We posit that this individual may have felt more confident about their familiarity or knowledge of these strategies beforehand, but then after attending the workshop they may have realized that they are not as familiar with them as they originally thought. However, it might also be due to simple day-to-day variation in their responses.

Respondents who indicated at least some familiarity with active learning were asked if they could explain what they know about active learning strategies and what might limit them from using active learning strategies in their role as a GTA. The responses to these questions are illustrated as word clouds in Figures 6.8 and 6.9. Figure 6.8 highlights that the word "students" is associated with active learning strategies which is not surprising since these strategies are utilized in a *student*-centred teaching approach. Figure 6.9 illustrates that time is one of the major barriers GTAs face when thinking about using active learning strategies in their roles.

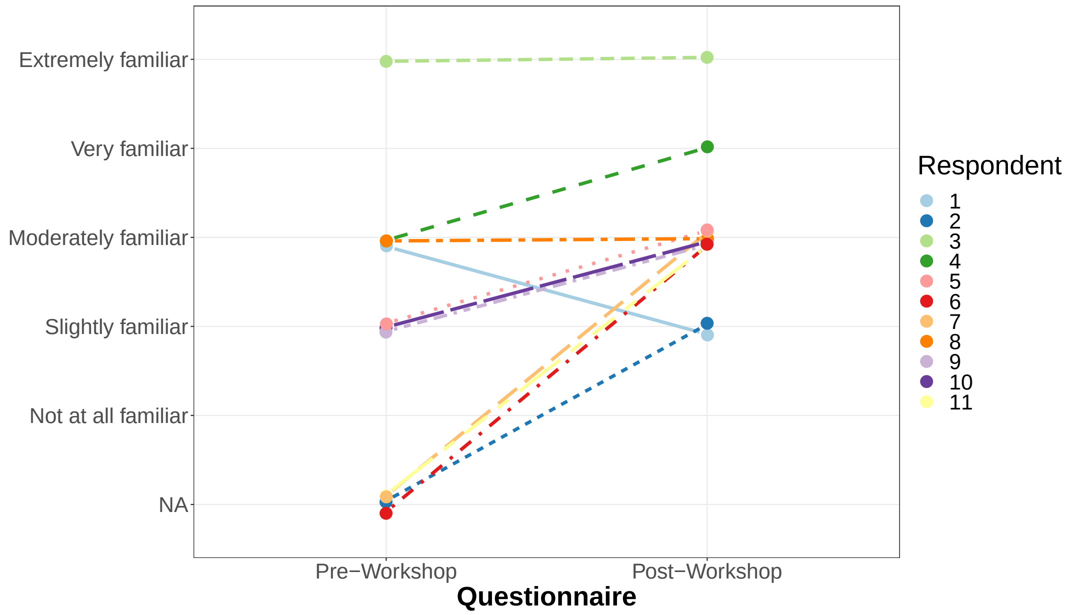


Figure 6.7: Connected scatterplot of responses to the question “Fill in the blank: I am (blank) with active learning strategies used in my discipline.” NA refers to no response.

Figure 6.8: Word cloud of responses to the question “What do you know about active learning strategies?”

Figure 6.9: Word cloud of responses to the question "What would limit you from using active learning strategies in your role as a GTA?"

Selected responses for the "What do you know about active learning strategies?" question follow:

> ***Pre-workshop***:
>
> ***Post-workshop***: *The definition and some examples.*
>
> —Ph.D. student, Mathematics

> ***Pre-workshop***: *Interactive*
>
> ***Post-workshop***: *Active learning is the learning environment where the learner participates actively in the generation of knowledge.*
>
> —Ph.D. student, Actuarial Science

> ***Pre-workshop***: *Let students ask questions actively and teach them course material by answering their questions, rather than teach them with all the course content.*

Post-workshop*: It is a student-centered learning strategy, giving students opportunities to show what they learn, which is helpful to let them know what they don't know.*

—Ph.D. student, Statistics

Selected responses for the "What would limit you from using active learning strategies in your role as a GTA?" question follow:

Pre-workshop*: Time and other duties.*

Post-workshop*: Activities other than GTA duties.*

—Ph.D. student, Applied Mathematics

Pre-workshop*: There are too many students in a course.*

Post-workshop*: It take a lot of time, and it requires the cooperation of students.*

—Ph.D. student, Statistics

Pre-workshop*: The burden of content coverage.*

Post-workshop*: The only limit for active learning strategies is our community's collective lack of knowledge outside of standard active learning strategies for math & statistics. The problem is probably being a pioneer and having to justify your teaching.*

—Master's student, Statistics

Pre-workshop*: We are told by professors/lecturers ... what to do, if we are given tutorial hours. It also may require more prep hours, for which we are not paid.*

> ***Post-workshop***: *We are generally not given much opportunity for teaching, even in tutorials, and when it does occur the professor usually has a set of problems they want us to cover. Since we are not in charge of the class we cannot make tutorials mandatory or adjust the marking scheme, so we may see loss of student interest in tutorials.*
>
> —Ph.D. student, Statistics

Figure 6.10 shows how respondents feel about their GTA roles and duties in general. We see that most of the responses stayed the same except for Respondents 5 and 10, who changed from feeling indifferent to very excited and moderately excited, respectively, after attending the workshop. Both of these respondents were seasoned GTAs (i.e., with over five academic terms of GTA experience) who previously attended various teaching-related professional development programs.

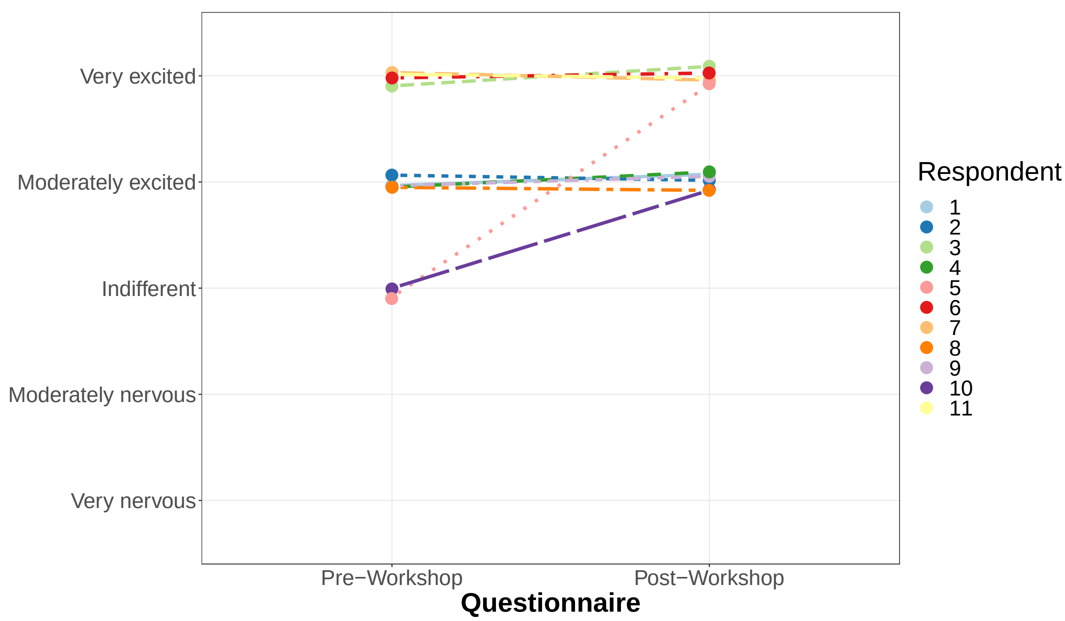


Figure 6.10: Connected scatterplot of responses to the question "Fill in the blank: I feel (blank) about my GTA roles and duties."

Figure 6.11 shows how comfortable the respondents were with incorporating active learning strategies in labs, tutorials, and office hours. We see that one respondent had a positive change (i.e., from not comfortable to comfortable), one respondent had a negative change, and four respondents remained unchanged. Interestingly, of the five respondents who did not provide a response to this question in the pre-workshop questionnaire, four of them responded that they felt comfortable incorporating active learning strategies in labs, tutorials, and/or office hours after attending the workshop. This could be a result of the workshop or from day-to-day variation in the responses. Overall, Respondents 2, 10, and 11 felt that they were still not comfortable incorporating such strategies even after attending the workshop due to their lack of experience, the time-consuming nature of active learning, and that they still did not have enough training to be able to develop active learning strategies for material in mathematics.

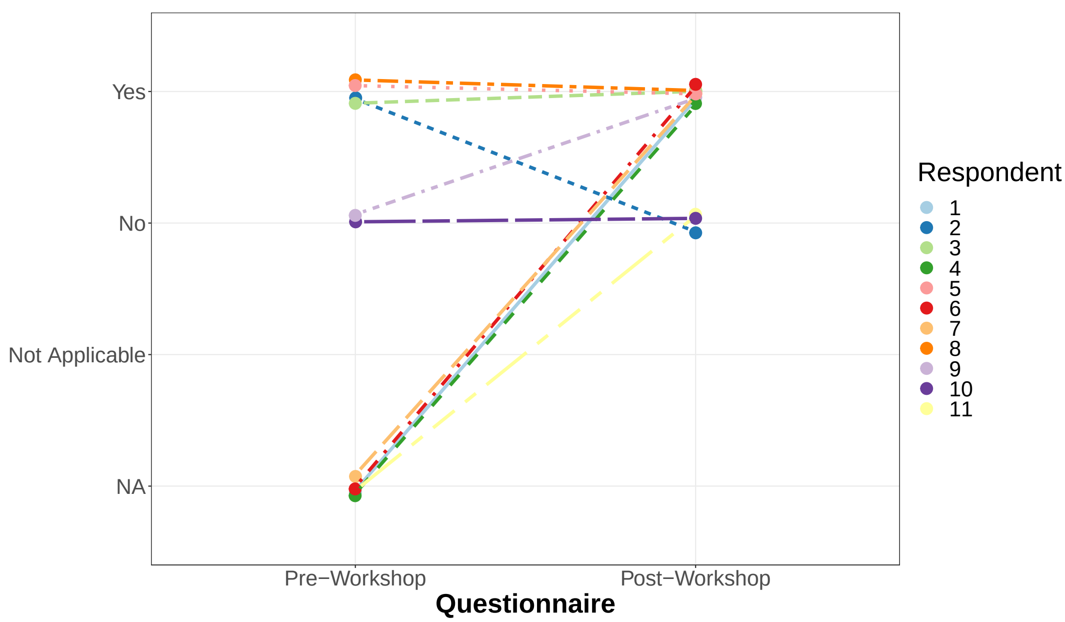


Figure 6.11: Connected scatterplot of responses to the question "As a GTA, I feel comfortable incorporating active learning strategies in labs, tutorials, and/or office hours." NA refers to no response.

If an individual responded "yes" to the previous question about their comfort level with incorporating active learning strategies, we asked them to provide an example of a strategy that they would use. If they responded "no", then we asked them to explain why they were not comfortable incorporating these activities. Selected responses to these questions are:

> ***Pre-workshop*** *("NA"):*
>
> ***Post-workshop*** *("yes"): For office hours I could divide students into teams based on which questions they have troubles with, and let students take initiative.*
>
> —Ph.D. student, Financial Modelling

> ***Pre-workshop*** *("no"): I would want to be properly trained (and compensated for that training time, as part of my GTA hours) so that I could execute the strategies effectively. I would also want the professor's approval, as I wouldn't want to go against their methods or cause a division between my section and those of the other GTAs.*
>
> ***Post-workshop*** *("yes"): I would use a more student-guided approach to questions I am asked in office hours, if we have an appropriate amount of time.*
>
> —Ph.D. student, Statistics

> ***Pre-workshop*** *("yes"): A trivially-implementable active-learning strategy for GTA-led labs, seminars, or tutorials would be Problem-Based Learning. You hone in on a particularly robust problem and solve it as a class. It allows for independent work, group work, teachable moments, discussion, and richer understanding of the course material.*
>
> ***Post-workshop*** *("yes"): I would use an entry ticket of a poll near the beginning of class to learn how well students are learning the content.*
>
> —Master's student, Statistics

The respondents developed interesting active learning strategies that they could implement in tutorials, labs, and office hours during the workshop, as mentioned in the quotes provided (e.g., student teamwork during office hours; polling questions to start a lab or tutorial). Many of the reasons why respondents did not feel comfortable incorporating active learning strategies were very similar to what would limit them from using such strategies (e.g., lack of pay, time constraints, etc.).

Figure 6.12 assesses respondents' interest in learning more about active learning strategies that they can use as GTAs in their respective disciplines. We see that eight respondents were, at best, as interested in learning more about active learning after the workshop and that three respondents were less interested in learning more.

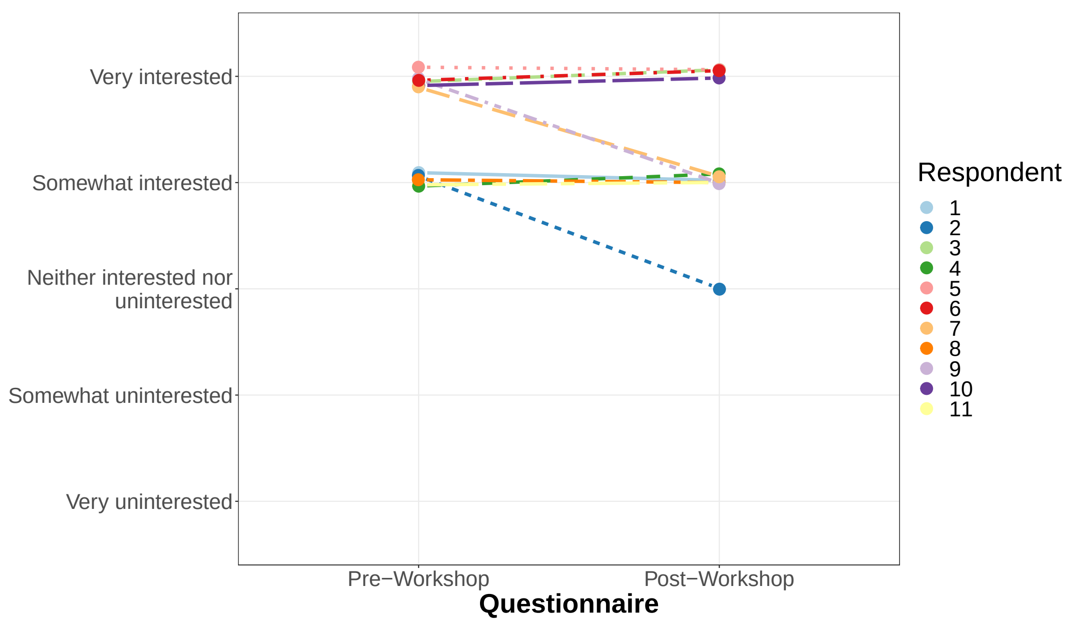


Figure 6.12: Connected scatterplot of responses to the question "Fill in the blank: I am (blank) in learning more about active learning strategies that I can use as a GTA in my discipline."

Figure 6.13 illustrates whether respondents think it is important or unimportant to use active learning strategies in undergraduate courses in their respective disciplines. Roughly half the respondents had the same response before and after the workshop while three respondents decreased and two respondents increased in their responses. Seven respondents felt that it was at least somewhat important to use active learning strategies in undergraduate courses after attending the workshop. Interestingly, we see more variability in the responses to this question because, although the literature explains the importance of active learning strategies in all disciplines, the disciplines of mathematics and/or statistics as a whole may be lagging in their incorporation or prioritization of active learning activities in undergraduate courses. The responses may also reflect the impression that students have about the importance of active learning influenced by their faculty, department, and/or supervisors.

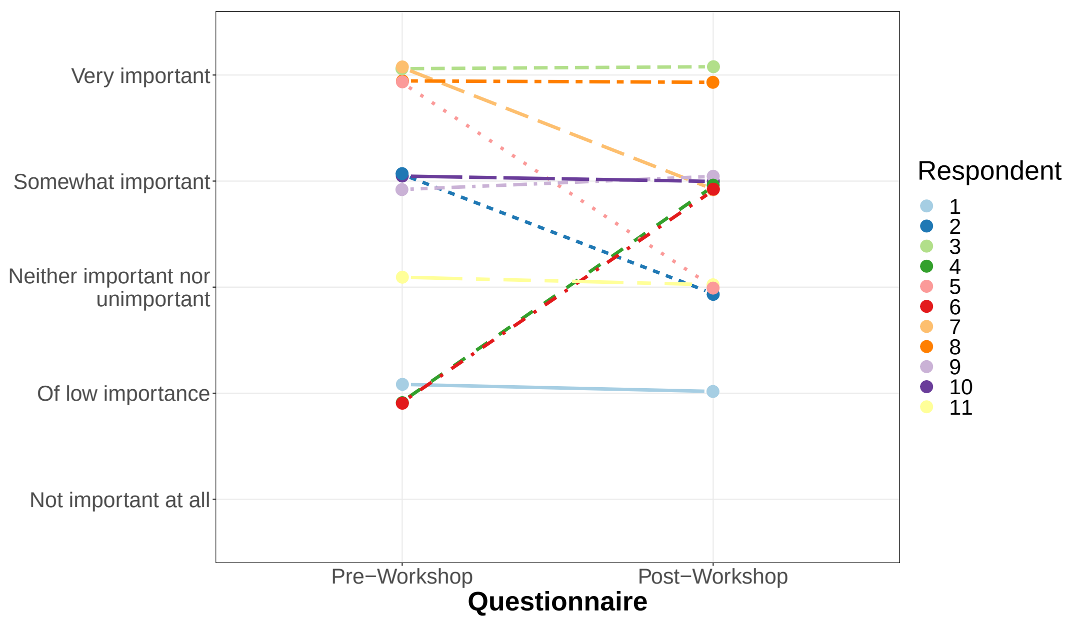


Figure 6.13: Connected scatterplot of responses to the question “Fill in the blank: In my discipline, it is (blank) to use active learning strategies in undergraduate courses.”

At the end of the workshop, we gave each participant a handout that summarized important ideas from the workshop and provided a list of references related to active learning (see Appendix D). We asked respondents in both questionnaires if they knew where to find active learning strategies that could be used in their discipline and the results are shown in Figure 6.14. Six respondents had an increase, three had a decrease, and two stayed the same in their responses. Most individuals who had an increase in their responses went from not knowing where to find resources on active learning or being neutral, to agreeing that they know where to find such resources. It is interesting to see that the first respondent went from agreeing to disagreeing — the workshop may have helped them realized that they, in fact, do not know where to look for resources even though they thought they had or that the workshop was lacking in terms of providing these resources.

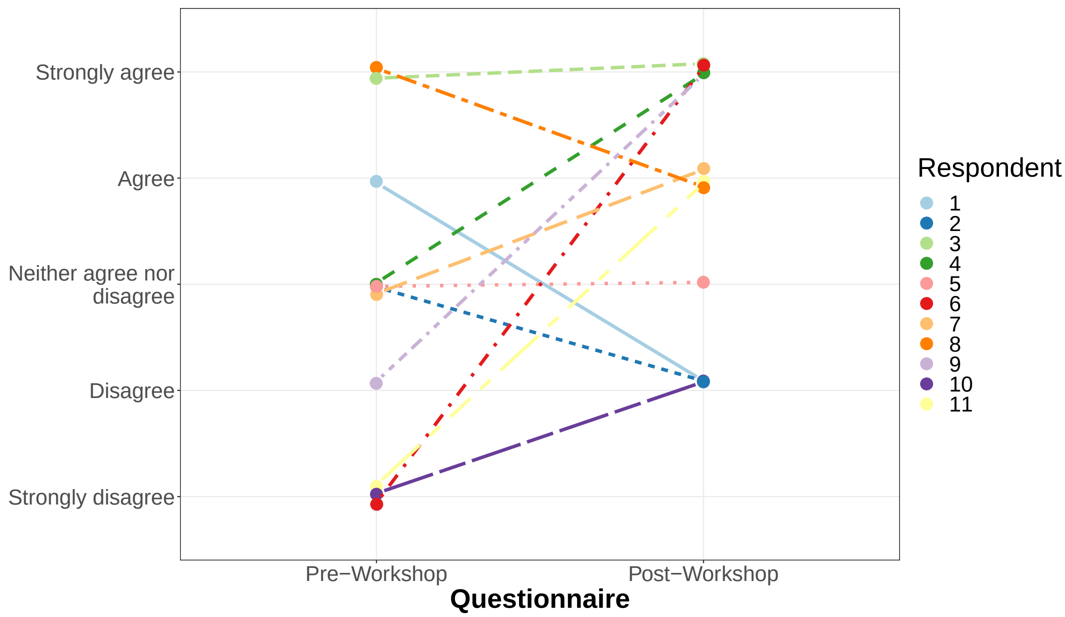


Figure 6.14: Connected scatterplot of responses to the question "I know where to find resources on active learning strategies that can be used in my discipline."

On the post-workshop questionnaire, we asked respondents two questions related to the workshop: how they felt about completing their GTA assignment after completing the workshop and what type of impact this workshop had on their perception of teaching as a GTA in their discipline. Table 6.1 shows the responses to those questions. The respondents either felt the same or more excited about their GTA assignment after attending the workshop and 91% of them felt that the workshop had a positive or very positive impact on their perception of teaching.

Table 6.1: Counts and percentages of responses to the last two questions on the post-workshop questionnaire.

Q: After completing this workshop, I feel (blank) about completing my GTA assignment.		**Q:** This workshop has had a (blank) impact on my perception of teaching as a GTA in my discipline.	
Response Categories	*Count (%)*	*Response Categories*	*Count (%)*
More excited	4 (36%)	Very Positive	2 (18%)
		Positive	8 (73%)
About the same	7 (64%)	Neither positive nor negative	1 (9%)
Less excited	0 (0%)	Negative	0 (0%)
		Very negative	0 (0%)

At the end of both questionnaires, we asked respondents to leave any comments related to the workshop, active learning strategies, or their experience as a GTA. The following comments were provided:

> *Thanks for running this workshop!*
>
> —Master's student, Statistics

> *It would be nice if more GTAs participate in these activities to share more experiences and perspectives.*
>
> —Ph.D. student, Statistics

> *I don't feel we got enough training for coming up with active learning strategies for material in math. This workshop mainly focused on teaching us active learning strategies through these strategies, which is cool and interesting, but*

> *not enough. In math, we have [a] lack of ideas of how to apply this to the material we have in the courses. Another workshop with practical examples for delivering Calculus, Linear Algebra, ODEs [Ordinary Differential Equations], etc. would make this training more complete and make us feel for comfortable with active learning strategies in our discipline.*
>
> —Ph.D. student, Applied Mathematics

> *I think for any meaningful implementation of active learning in tutorials/labs it has to come from the top down (i.e., from professors). GTAships do not typically include enough allotted time for preparation of tutorials, and generally attendance is quite poor for introductory courses. It should not be expected of GTA's to go over their hours, especially when there is generally little to no recognition of good or bad GTA work in the department, and no possible reward (such as raises, promotions) simply due to how the courses are structured. The incentive is very low.*
>
> —Ph.D. student, Statistics

6.4 Discussion

It is clear from the literature that both short and long teaching development programs have several benefits for GTAs. Fostering the growth of the teaching ability of GTAs can also enhance their abilities to communicate their research and help them become better researchers in general since many of the communication and other skills they develop are transferable. Therefore, training programs on effective teaching and pedagogy add to the overall professional development of graduate students and should boost their credibility as a researcher. Teaching-related workshops can also serve as a gateway to further teaching development since they often help participants understand the benefits that can be gained from them. However, GTA teaching-related training and development programs within statistics and mathematics often lack pedagogical insights that can strengthen the teaching styles used by GTAs and even instructors.

We developed this pre-post study to investigate how participating in a discipline-specific teaching development program on active learning for GTAs in mathematics and statistics impacts their perceptions of teaching. A pre-post self-report survey design (Salkind, 2010) is the most widely used design where participants are asked questions before an intervention (pretest), participate in the intervention, and are again asked questions after the intervention (posttest). This design measures changes in participant knowledge or attitudes regarding the intervention content. It is viewed as a rigorous method that provides credible results and measures the same person at two time intervals, reducing many sources of bias. It is important to note that pre-post surveys only assess the *respondents' perceptions* of their learning. Thus, interpretations of the differences between pre and post ratings are limited to what respondents *think* they learned or how much they *think* they changed.

In general, it appears that the workshop helped participants to be able to define active learning and provided them with active learning examples. Most respondents noted that time constraints, other duties (GTA and non-GTA), class sizes, space, increased preparation time, lack of pay for preparation time, and a lack of autonomy when running tutorials/labs/office hours would limit them from using active learning strategies in their GTA roles. These concerns are valid and have been well documented as obstacles that instructors may face, or believe that that may face, when implementing active learning activities in courses (Faust and Paulson, 1998; Braun et al., 2017). If instructors and professors have these concerns, then it is understandable that GTAs — who often have less teaching experience, fewer opportunities for teaching professional development, and may receive little-to-no training on teaching or active learning — have them as well.

The results from the responses were mixed. Furthermore, given the small number of participants, the discussion of results that follows is a qualitiative description rather than a quantitative analysis. For respondents who did change their responses after attending the workshop, we found that the workshop may have: slightly encouraged a few respondents to consider using active learning strategies, helped some respondents become slightly more familiar with active learning strategies, and increased over half of the respondents' knowledge of where to find resources on active learning strategies that could be used in their discipline. Additionally, for respondents who did change their responses after attending the workshop, we found that the workshop may have: slightly discouraged a few respondents' interest in learning more about active learning strategies, and both increased and decreased some respondents' perception towards the importance of active learning strategies used in undergraduate courses in their disciplines. However, Table 6.1 illustrates that the workshop did not appear to negatively impact respondents' excitement for completing their GTA assignments nor did it negatively impact their perception of teaching as a GTA — both of which are good.

There are some limitations to the study design that we must consider. The pre-post design is known to have flaws like the response shift bias (Howard, 1980) where participants' framework of understanding a question would shift between the pre and post periods, resulting in inaccurate assessments of their pre-program knowledge due to their lack of understanding at that time which can underestimate the program effect (Skeff et al., 1992). An alternative design was proposed in the late 1970's — called the post-then-pre design (Colosi and Dunifon, 2006; Kanevsky, 2016) — to reduce or eliminate the response shift bias. Additionally, it is a more convenient design since both measures are taken at the same time, making it less burdensome and time-consuming for participants. The limitations of the post-then-pre design include bias when recalling memories for the "pre" section, even with short time periods between pre and post, and participants reporting improvement (even subconsciously) to justify the time and energy they invested in program attendance.

Self-reporting surveys, such as ours, are always vulnerable to bias since participants may only answer what they think the evaluator wants. Thus the inaccuracy of greatly differing self-assessments may not provide a reliable measure. Additionally, self-selection bias is an issue in our study design since participants self-selected to participate in the study and attend the workshop. The main issue is that our workshop participants were most likely GTAs who were already interested in learning more about teaching and teaching methods that could assist them. This may be assessed in future studies by including questions asking why the participant attended the workshop. Consequently, it is reasonable to infer that our sampled population did not encapsulate our target population of all GTAs within the University of Western Ontario's School of Mathematics and Statistics, including those who are interested in teaching professional development and those who are not.

A major limitation of our study is the low participation rate of 12 individuals of which 11 provided responses to the pre-post questionnaires. This hampered quantitative data analysis collected from the questionnaires to investigate whether the intervention (i.e., workshop) had an impact on participants' perception of teaching via large sample asymptotics. However, nonparametric tests like the exact Binomial test or Fisher's exact test, where responses would need to be regrouped into dichotomous levels (e.g., strongly agree and agree into agree), would be an option for analyzing small sample sizes. These analyses were not performed as we considered these responses as part of a pilot study, with the intention to extend the study period to offer a second iteration of the workshop sessions where we will be able to conduct an analysis with reasonable power to detect changes. We note that another option for analyzing the quantitative data in future includes fitting ordinal logistic regression models (Harrell, 2015) that use the demographic data as predictors, incorporating fixed and possibly interaction terms to account for effects of the different workshop sessions that respondents attend. Regardless of the methods employed, each question requires appropriate thought and consideration in terms of how to analyze the data since different questions utilize positive versus negative framing and unipolar versus bipolar items.

Possible solutions for increasing the participation rate of the workshop may include greater advertisement, offering different mediums such as an online version to enhance flexibility, or offering the workshop earlier in the fall term before GTAs get busy with their own courses, research, and GTAships. Recall that the workshops were offered starting in mid-September which might have already been "too late" in the term. Also, we must be cognizant of the general timing since September 2021 was the first in-person month at Western University after 18 months of online living due to the COVID-19 pandemic. Concerns about COVID-19, returning to in-person activities, and quarantines for certain graduate students[4] may have potentially kept participation in our in-person workshop low.

The next iteration of our workshop and study will most likely benefit from the following adjustments:

1. Having the Department, School, or Faculty promote the workshop to its GTAs;
2. Offering fewer sessions (only one or two) earlier in term during the first or second week of classes before GTAs become too busy with their own workloads;
3. Switching to a post-then-pre survey design instead of a pre-post design; and
4. Administering paper questionnaires with time for completion built into the workshop so that participants can respond before they leave the room.

It may only be necessary to implement some and not all of the noted adjustments. We believe that collaborating with the respective Department, School, or Faculty to offer this workshop would have the most significant impact on our participation and response rates, particularly in regard to the workshop promotion. Promoting the workshop would

[4] As per Western University's COVID policy, students returning to campus from countries outside of Canada were required to quarantine for a 14-day period.

emphasize the value that faculty members place on active learning techniques. It would also demonstrate that the Department, School, or Faculty prioritizes the teaching professional development of their graduate students and the many benefits that this type of training offers early career individuals — a crucial element of graduate students' academic journey that is frequently overlooked.

Pentecost et al. (2012) developed and analyzed a specialized GTA training program for graduate students that ran for four years and permanent department faculty and staff assumed the leadership roles in planning and leading the sessions alongside the project staff for the last two years. The shift of faculty and staff becoming increasingly involved in taking over the training program ensured its continued endurance and growth as a departmental asset. Our long-term goals for this research consist of (1) repeating this study to provide concrete evidence that discipline-specific training programs for GTAs in mathematics and statistics are important and necessary for their overall career development, and (2) refining the workshop to create a sustainable, high-quality GTA teaching-related training and development program that could be integrated into all mathematics and statistics departments across Canadian universities.

Chapter 7

Conclusions and Future Work

The first part of this dissertation used the principles of interdisciplinary knowledge exchange to develop novel techniques for the study of wildland fire lifetimes. The detection and action phases of fully suppressed wildland fires in the Sioux Lookout District of northwestern Ontario were investigated by fitting multi-state models to determine which factors influence the time spent in each phase and how the phases are interrelated. Williams et al. (2017) noted that if a previous phase, used as a predictor, is found to have evidence of an effect on the current phase, then there is evidence to suggest that the Markov property does not hold, indicating that a semi-Markov model is more appropriate. Our work shows that longer detection phases are associated with longer action phases, implying that the semi-Markov models such as the clock-reset multi-state model or the joint frailty models are more appropriate when modelling wildland fire lifetimes. Although we were able to determine the influential factors, the results from the multi-state models lacked insightful information about the relationship between the two phases. Joint frailty models were employed to allow for correlation between the detection and action phases by incorporating fire-specific random effects (or *frailties*) in the models. We found that the action phase lengths may be increasing over time.

Comparisons of models with different frailty distributions, frailty forms, and baseline hazard functions determined the preferred models for both the human- and lightning-caused fires. Although different frailty distributions and baseline hazard functions were identified by the model selection techniques, both models agreed on the factor loading and independent (FLI) form of the frailty term allowing for both the connection between the two phases and the flexibility for each phase to differ. Results from a simulation study highlighted the FLI form as a dominant model form when ranking using two Bayesian-based ranking methods (WAIC and PSIS-LOO). A communiqué is being developed in collaboration with a member of the Ontario Ministry's Aviation, Forest Fire and Emergency Services. This deliverable is a tangible component of the collaborative knowledge exchange that was employed for this work. It will provide the Ministry and its fire management practitioners with a broad overview of the results and insights gleaned from investigating these wildland fire lifetime phases and to offer areas of potential further investigation.

The next step in this research involves utilizing more flexible baseline hazard functions like the piecewise exponential with optimal allocation for join point placements or cubic B-splines as mentioned in Nathoo and Dean (2008) to add extra flexibility in the joint frailty modelling framework. Another step would be to examine additional forms of the frailty term, such as the multivariate form from Xi et al. (2020). Using a multivariate form for the fire-specific random effects where the frailties follow a multivariate normal distribution with a nonzero covariance allows the lifetime phases to be interrelated in a different way than using the shared factor loading parameter. Additionally, the modelling framework can be extended by splitting the fire lifetime into further phases. For instance, the action phase is comprised of the dispatch (report to getaway), travel (getaway to initial attack), and suppression (attack to under control) phases. By investigating the fire lifetime at smaller intervals we will be able to better understand the relationships that these phases have with one another and how previous phases affect subsequent ones.

The fire weather variables previously discussed in Section 3.3.2 that were considered as predictors in the modelling are the observed values on the day a fire was ignited. This is a limitation of our analysis, especially for fires that lasted several days or more. We can improve on the use of a single value for a variable that is determined by the changing weather when modelling wildland fires over time. Incorporating time-varying predictors into future models where the predictors are "reset" at the start of each phase addresses this issue. However, rather than simply using the historical weather data, forecasts of fire weather variables could also be used as predictors. A comparison of the prediction results obtained from using actual and forecasted time-varying fire weather variables in the models should be performed to investigate any changes of the variability in the model estimates. For instance, future researchers should fit joint frailty models to training data using the actual (observed) time-varying weather data and make predictions on testing data. Then they should make predictions for the same models on the testing data using the forecasted time-varying weather data instead of the actual data. The differences in the predictions would be compared, since they are attributable to any error in the weather predictions, while also accounting for any uncertainty in the predictions. A case study of fewer, but meaningful, wildland fires for each model may be necessary since both the daily and forecasted (1 day, 2 day, 4 day) weather data would be required for every fire the dataset. Meaningful fires would be carefully chosen in collaboration with the Forest Fire Science Specialist from the Ontario Ministry and future researchers must ensure that the study has enough fires to make appropriate inferences.

This research may also be extended to consider the spatial context more explicitly. Past research by Morin et al. (2019) fit frailty models for the control time of wildland fires in the former Intensive Fire Management Zone in Ontario. Their objective was to investigate spatial differences in their study area by utilizing Cox PH shared frailty models with a Gaussian random effect to modify the hazard for fires within each spatial partition. Ontario's Sioux Lookout District can be spatially partitioned based on which fire management zone (FMZ) fires are located in. Prior to 2004, the fire region was partitioned into three FMZs, receiving different levels of protection. Fires in the former Intensive Fire Management Zone were suppressed as soon as resources were available whereas fires in the former Extensive Fire Management Zone were monitored and most were left to burn out so long as they did not threaten communities or other important values. After 2004, the province of Ontario was divided into six FMZs (OMNR, 2004) based on common management objectives, land use, fireload, and forest ecology. Each zone had its own management objectives and fire response direction. The zones changed again in 2014 when the Ministry moved to their framework of "appropriate response" (OMNRF, 2014b). As a first step, our joint frailty models can be extended by also incorporating a FMZ Gaussian spatial random effect with the understanding that the spatial effect will be constrained to the FMZ partitions. Ultimately, fitting the joint frailty models with a Gaussian random field using the Integrated Nested Laplace Approximation (INLA) approach (Rue et al., 2009), computed with the `R-INLA` package (Lindgren and Rue, 2015), would add more spatial flexibility. These spatial extensions require careful consideration since the fire management strategy changed twice over the study period, affecting the zones that partitioned the Sioux Lookout District.

The second component of this dissertation consisted of the development and effectiveness of a training program for graduate teaching assistants in the mathematical and statistical sciences. The workshop focussed on active learning techniques — techniques that aid in the exchange of knowledge between students and instructors (or GTAs) — along with their benefits, limitations, and examples that GTAs could draw on in a variety of learning environments (e.g., tutorials, labs, office hours, etc.). We performed a survey study where participants were asked to attend the workshop and participate in pre-post survey questionnaires to assess the workshop and answer the research question: **How does participation in a discipline-specific teaching development program on active learning for graduate teaching assistants in mathematics and statistics, offered by their School of Mathematical and Statistical Sciences, impact their**

perceptions of teaching?

We found that the workshop helped participants define active learning, provided them with knowledge of where to find resources on active learning strategies, and did not have a negative impact on their perceptions of teaching as a graduate teaching assistant. In future, we plan to refine the structure of the workshop, collaborate with a Department, School, or Faculty to offer and promote the workshop, and perform the study again with the aim of a higher participation rate. The long-term goal of this work is to create a sustainable, high-quality GTA teaching-related workshop that can be integrated into the GTA training and development programs offered at other mathematics and statistics departments across Canadian universities.

Appendix A

Chapter 3 Supplementary Material

A.1 Converting Data from Wide Format to Long Format

The data for both the human- and lightning-caused fires are converted from wide format to long format. The first six rows of the long format dataset for the human-caused fires are shown in Table A.1.

Table A.1: First six rows of long format dataset for human-caused Sioux Lookout fires.

id	from	to	trans	Tstart	Tstop	time	status
1	1	2	1	0	0.574	0.574	1
1	2	3	2	0.574	0.625	0.051	1
2	1	2	1	0	0.049	0.049	1
2	2	3	2	0.049	0.132	0.083	1
3	1	2	1	0	0.042	0.042	1
3	2	3	2	0.042	0.056	0.014	1

The ‘id’ variable identifies the fire so the first two rows of data correspond to the same fire. The ‘trans’ variable identifies which transition has occurred and the ‘from’ and ‘to’ variables identify the starting and ending states of that transition. ‘Tstart’ and ‘Tstop’ correspond to the starting and stopping time (in days) of the fire for each phase (i.e., clock-forward time) whereas 'time' indicates the overall time spent in each phase (i.e., clock-reset time). The 'status' indicator variable represents whether an event was observed (status = 1) or censored (status = 0). The predictors associated with each phase were left out from the dataset. We see that all three of the fires were declared under control in less than 24 hours but that it took a much longer time to detect the first fire.

There are a total of 1,572 and 2,540 observations for the respective human- and lightning-caused fires after converting the data to long format (i.e., the number of observed fires multiplied by the two lifetime phases).

A.2 Diagnostic Plots

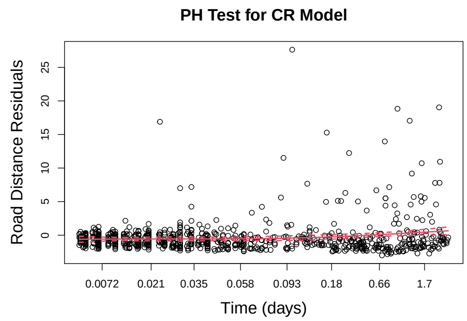


Figure A.1: Plot of the scaled Schoenfeld residuals for the road distance predictor from the human-caused CR multi-state model against time in days, along with a smooth curve of the residuals with 95% confidence bands.

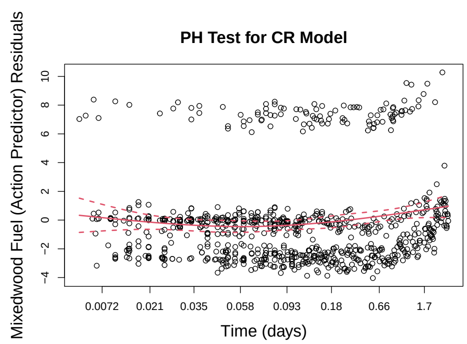


Figure A.2: Plot of the scaled Schoenfeld residuals for the mixedwood fuel predictor from the human-caused CR multi-state model against time in days, along with a smooth curve of the residuals with 95% confidence bands.

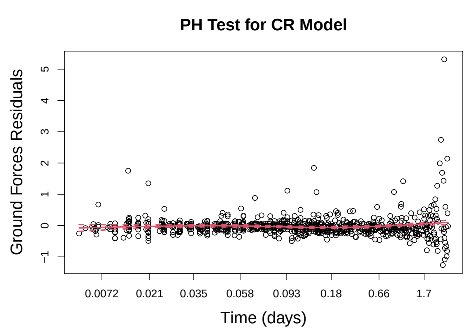


Figure A.3: Plot of the scaled Schoenfeld residuals for the ground forces predictor from the human-caused CR multi-state model against time in days, along with a smooth curve of the residuals with 95% confidence bands.

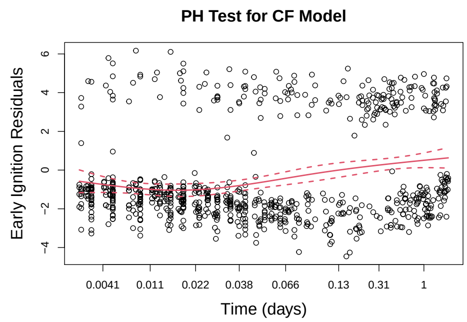


Figure A.4: Plot of the scaled Schoenfeld residuals for the early ignition predictor from the human-caused CF multi-state model against time in days, along with a smooth curve of the residuals with 95% confidence bands.

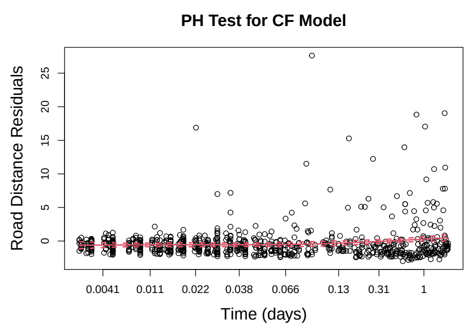


Figure A.5: Plot of the scaled Schoenfeld residuals for the road distance predictor from the human-caused CF multi-state model against time in days, along with a smooth curve of the residuals with 95% confidence bands.

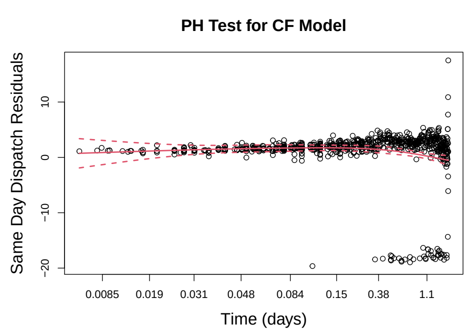


Figure A.6: Plot of the scaled Schoenfeld residuals for the same day dispatch predictor from the human-caused CF multi-state model against time in days, along with a smooth curve of the residuals with 95% confidence bands.

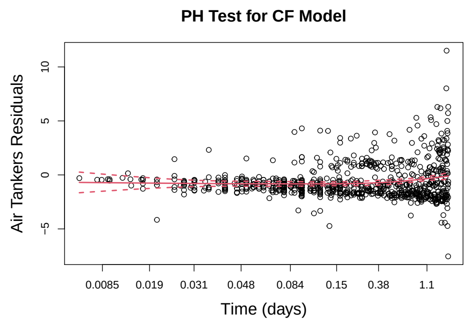


Figure A.7: Plot of the scaled Schoenfeld residuals for the air tankers predictor from the human-caused CF multi-state model against time in days, along with a smooth curve of the residuals with 95% confidence bands.

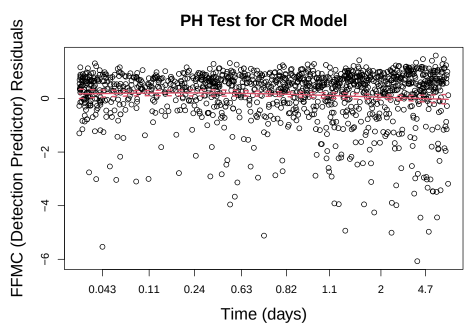


Figure A.8: Plot of the scaled Schoenfeld residuals for the FFMC (detection) predictor from the lightning-caused CR multi-state model against time in days, along with a smooth curve of the residuals with 95% confidence bands.

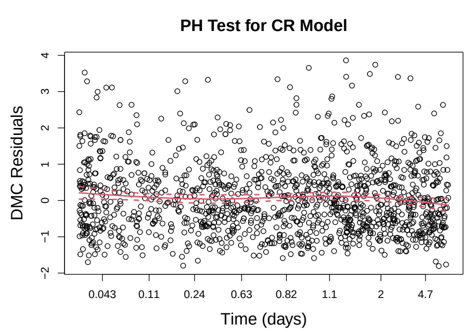


Figure A.9: Plot of the scaled Schoenfeld residuals for the DMC predictor from the lightning-caused CR multi-state model against time in days, along with a smooth curve of the residuals with 95% confidence bands.

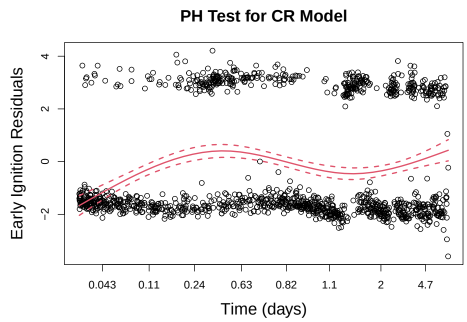


Figure A.10: Plot of the scaled Schoenfeld residuals for the early ignition predictor from the lightning-caused CR multi-state model against time in days, along with a smooth curve of the residuals with 95% confidence bands.

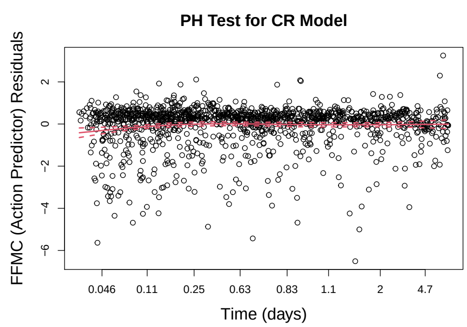


Figure A.11: Plot of the scaled Schoenfeld residuals for the FFMC (action) predictor from the lightning-caused CR multi-state model against time in days, along with a smooth curve of the residuals with 95% confidence bands.

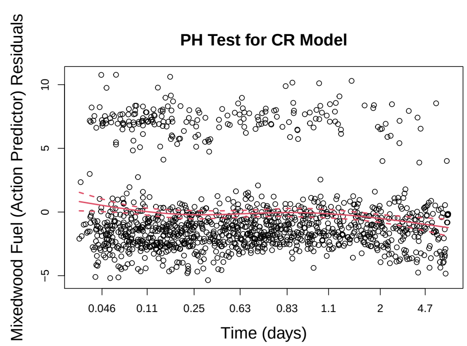


Figure A.12: Plot of the scaled Schoenfeld residuals for the mixedwood fuel predictor from the lightning-caused CR multi-state model against time in days, along with a smooth curve of the residuals with 95% confidence bands.

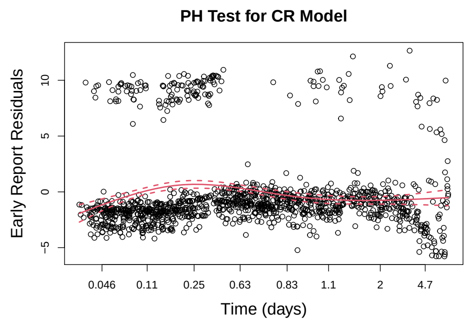


Figure A.13: Plot of the scaled Schoenfeld residuals for the early report predictor from the lightning-caused CR multi-state model against time in days, along with a smooth curve of the residuals with 95% confidence bands.

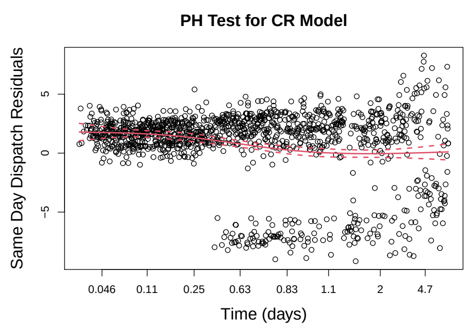


Figure A.14: Plot of the scaled Schoenfeld residuals for the same day detection predictor from the lightning-caused CR multi-state model against time in days, along with a smooth curve of the residuals with 95% confidence bands.

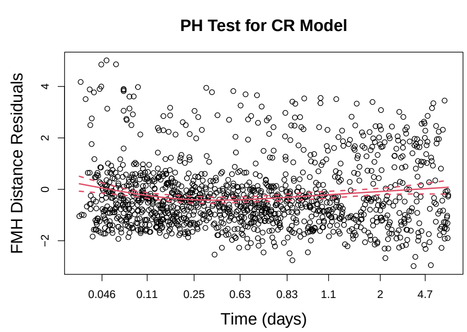


Figure A.15: Plot of the scaled Schoenfeld residuals for the FMH distance predictor from the lightning-caused CR multi-state model against time in days, along with a smooth curve of the residuals with 95% confidence bands.

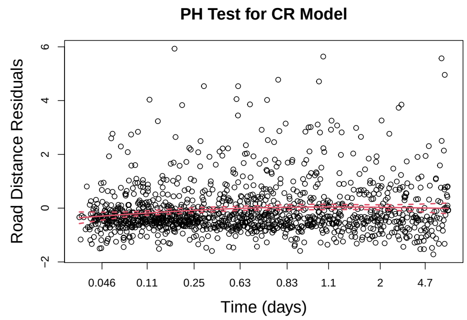


Figure A.16: Plot of the scaled Schoenfeld residuals for the road distance predictor from the lightning-caused CR multi-state model against time in days, along with a smooth curve of the residuals with 95% confidence bands.

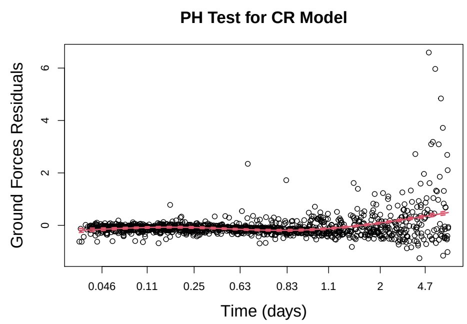


Figure A.17: Plot of the scaled Schoenfeld residuals for the ground forces predictor from the lightning-caused CR multi-state model against time in days, along with a smooth curve of the residuals with 95% confidence bands.

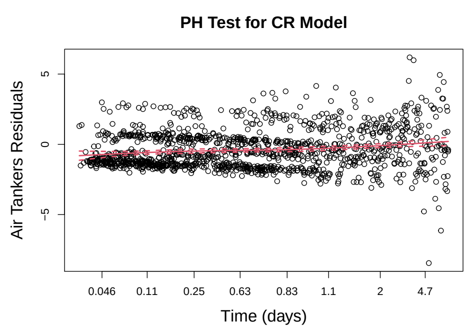


Figure A.18: Plot of the scaled Schoenfeld residuals for the air tankers predictor from the lightning-caused CR multi-state model against time in days, along with a smooth curve of the residuals with 95% confidence bands.

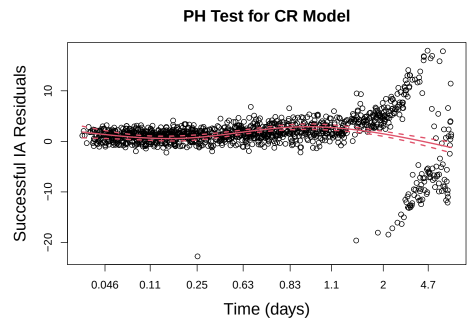


Figure A.19: Plot of the scaled Schoenfeld residuals for the successful IA predictor from the lightning-caused CR multi-state model against time in days, along with a smooth curve of the residuals with 95% confidence bands.

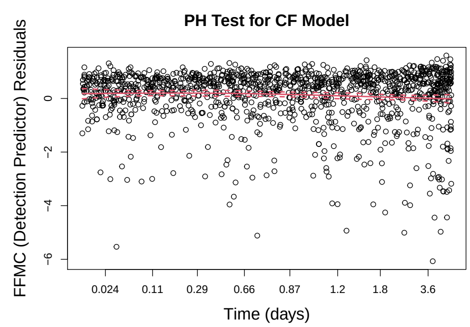


Figure A.20: Plot of the scaled Schoenfeld residuals for the FFMC (detection) predictor from the lightning-caused CF multi-state model against time in days, along with a smooth curve of the residuals with 95% confidence bands.

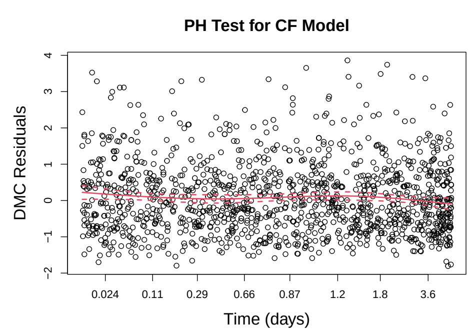


Figure A.21: Plot of the scaled Schoenfeld residuals for the DMC predictor from the lightning-caused CF multi-state model against time in days, along with a smooth curve of the residuals with 95% confidence bands.

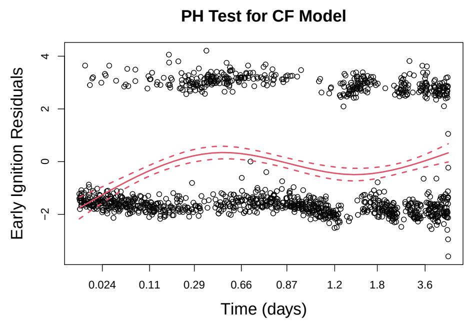


Figure A.22: Plot of the scaled Schoenfeld residuals for the early ignition predictor from the lightning-caused CF multi-state model against time in days, along with a smooth curve of the residuals with 95% confidence bands.

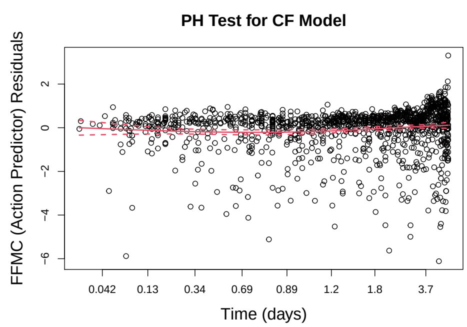


Figure A.23: Plot of the scaled Schoenfeld residuals for the FFMC (action) predictor from the lightning-caused CF multi-state model against time in days, along with a smooth curve of the residuals with 95% confidence bands.

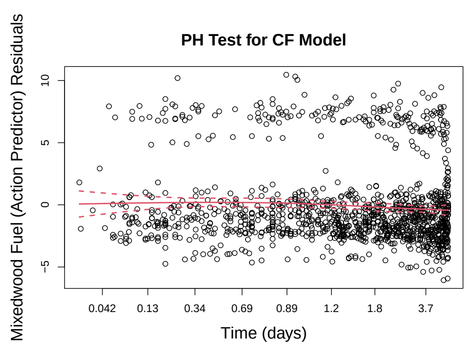


Figure A.24: Plot of the scaled Schoenfeld residuals for the mixedwood fuel predictor from the lightning-caused CF multi-state model against time in days, along with a smooth curve of the residuals with 95% confidence bands.

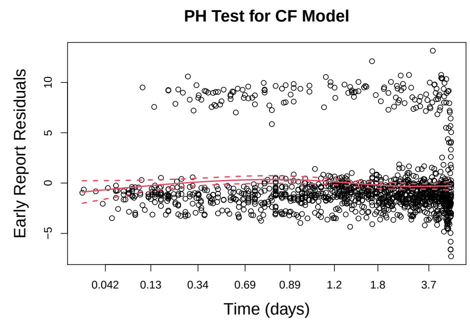


Figure A.25: Plot of the scaled Schoenfeld residuals for the early report predictor from the lightning-caused CF multi-state model against time in days, along with a smooth curve of the residuals with 95% confidence bands.

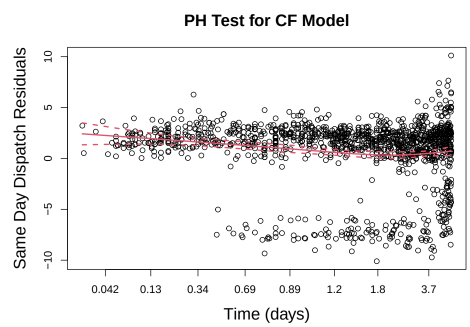


Figure A.26: Plot of the scaled Schoenfeld residuals for the same day dispatch predictor from the lightning-caused CF multi-state model against time in days, along with a smooth curve of the residuals with 95% confidence bands.

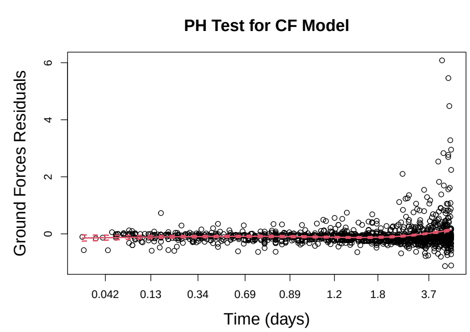


Figure A.27: Plot of the scaled Schoenfeld residuals for the ground forces predictor from the lightning-caused CF multi-state model against time in days, along with a smooth curve of the residuals with 95% confidence bands.

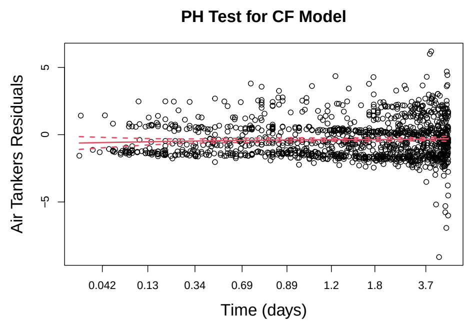


Figure A.28: Plot of the scaled Schoenfeld residuals for the air tankers predictor from the lightning-caused CF multi-state model against time in days, along with a smooth curve of the residuals with 95% confidence bands.

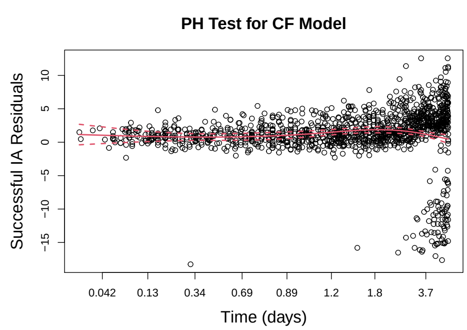


Figure A.29: Plot of the scaled Schoenfeld residuals for the successful IA predictor from the lightning-caused CF multi-state model against time in days, along with a smooth curve of the residuals with 95% confidence bands.

A.3 Estimated Transition Probability Plots

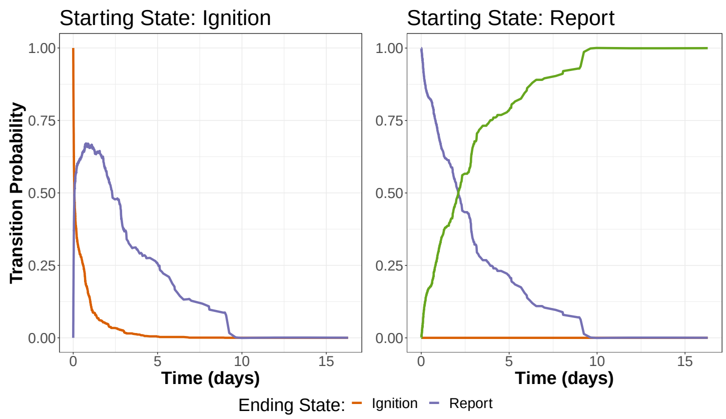


Figure A.30: Plots of the Aalen-Johansen transition probability estimates for the human-caused clock-forward multi-state model. The left plot shows the $\widehat{P}_{1l}(0,t)$ curves out of the ignition state and the right plot shows the $\widehat{P}_{2l}(0,t)$ curves out of the report state.

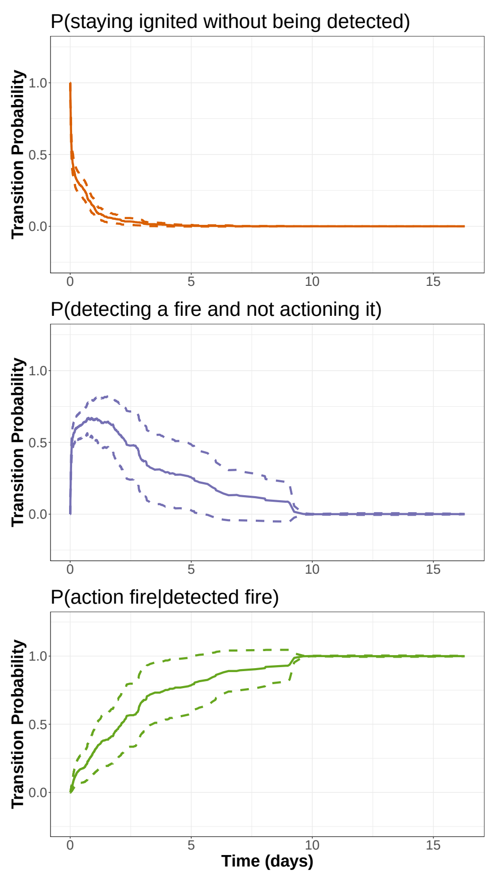


Figure A.31: Plots of specific Aalen-Johansen transition probabilities taken from Figure A.30 with the 95% confidence bands.

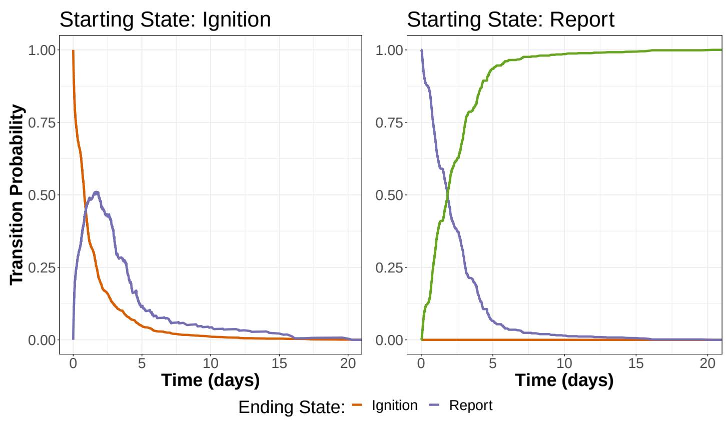


Figure A.32: Plots of the Aalen-Johansen transition probability estimates for the lightning-caused clock-reset multi-state model. The left plot shows the $\widehat{P}_{1l}(0,t)$ curves out of the ignition state and the right plot shows the $\widehat{P}_{2l}(0,t)$ curves out of the report state.

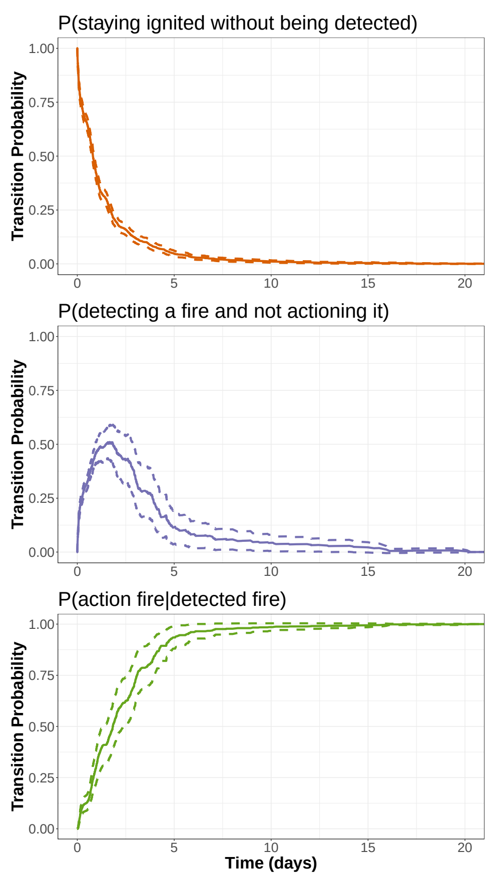


Figure A.33: Plots of specific Aalen-Johansen transition probabilities taken from Figure A.32 with the 95% confidence bands.

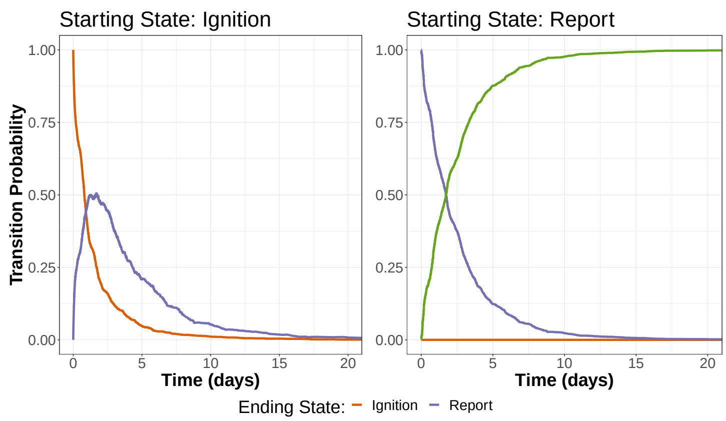


Figure A.34: Plots of the Aalen-Johansen transition probability estimates for the lightning-caused clock-forward multi-state model. The left plot shows the $\widehat{P}_{1l}(0,t)$ curves out of the ignition state and the right plot shows the $\widehat{P}_{2l}(0,t)$ curves out of the report state.

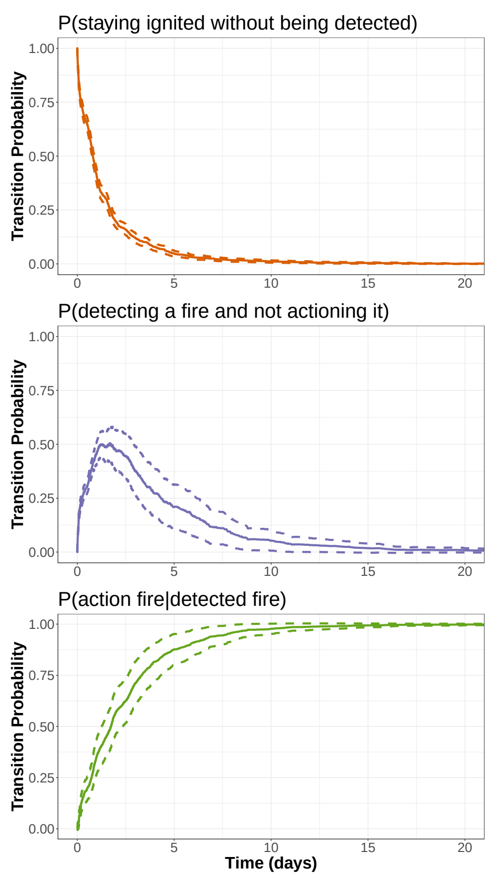


Figure A.35: Plots of specific Aalen-Johansen transition probabilities taken from Figure A.34 with the 95% confidence bands.

Appendix B

Chapter 4 Supplementary Material

B.1 Important Functions using the Weibull Baseline Hazard

A parametric Weibull baseline hazard function has the form $h_{0j}(t_{ij}|\mathcal{H}(t_{ij}^{-})) = \lambda_j \rho_j t_{ij}^{\rho_j - 1}$, where $\lambda_j > 0$ is the scale parameter and $\rho_j > 0$ is the shape parameter. The *hazard function* is

$$
\begin{aligned}
h_{ij}(t_{ij}|\mathcal{H}(t_{ij}^{-})) &= h_{0j}(t_{ij}|\mathcal{H}(t_{ij}^{-}))u_{ij}\exp(\boldsymbol{x}_{ij}^{'}\boldsymbol{\beta}_j) \\
&= \lambda_j \rho_j t_{ij}^{\rho_j - 1} u_{ij} \exp(\boldsymbol{x}_{ij}^{'}\boldsymbol{\beta}_j).
\end{aligned}
$$

The *survival function* is

$$
\begin{aligned}
S_{ij}(t_{ij}|\mathcal{H}(t_{ij}^{-})) &= \exp\left[-\int_0^{t_{ij}} \lambda_j \rho_j s_{ij}^{\rho_j - 1} u_{ij} \exp(\boldsymbol{x}_{ij}^{'}\boldsymbol{\beta}_j) ds_{ij}\right] \\
&= \exp\left[-\lambda_j \rho_j u_{ij} \exp(\boldsymbol{x}_{ij}^{'}\boldsymbol{\beta}_j) \int_0^{t_{ij}} s_{ij}^{\rho_j - 1} ds_{ij}\right] \\
&= \exp\left[-\lambda_j \cancel{\rho_j} u_{ij} \exp(\boldsymbol{x}_{ij}^{'}\boldsymbol{\beta}_j) \times \frac{t_{ij}^{\rho_j}}{\cancel{\rho_j}}\right] \\
&= \exp\left[-\lambda_j t_{ij}^{\rho_j} u_{ij} \exp(\boldsymbol{x}_{ij}^{'}\boldsymbol{\beta}_j)\right].
\end{aligned}
$$

Then the *density function* is

$$\begin{aligned} f_{ij}(t_{ij}|\mathcal{H}(t_{ij}^-)) &= h_{ij}(t_{ij}|\mathcal{H}(t_{ij}^-)) \times S_{ij}(t_{ij}|\mathcal{H}(t_{ij}^-)) \\ &= \lambda_j \rho_j t_{ij}^{\rho_j - 1} u_{ij} \exp(\boldsymbol{x}_{ij}^{'} \boldsymbol{\beta}_j) \exp\left[-\lambda_j t_{ij}^{\rho_j} u_{ij} \exp(\boldsymbol{x}_{ij}^{'} \boldsymbol{\beta}_j)\right], \end{aligned}$$

which results in $T_{ij} \sim \text{Weibull}\left(\lambda_j u_{ij} \exp(\boldsymbol{x}_{ij}^{'} \boldsymbol{\beta}_j), \rho_j\right)$.

B.2 Ranked Figures of ELPD Estimates

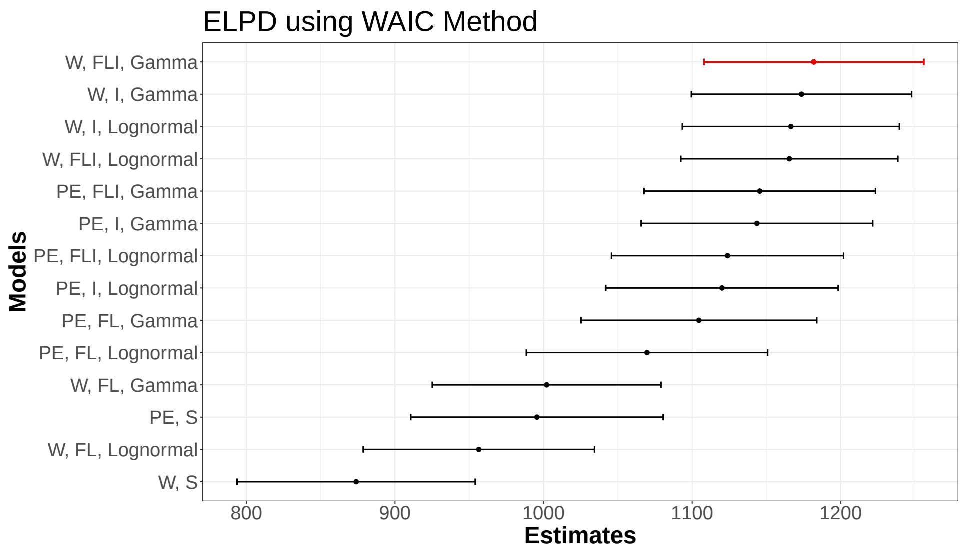


Figure B.1: Ranked comparisons of ELPD point estimates and standard errors, using the WAIC method, for human-caused wildland fires. The preferred model is highlighted in red.

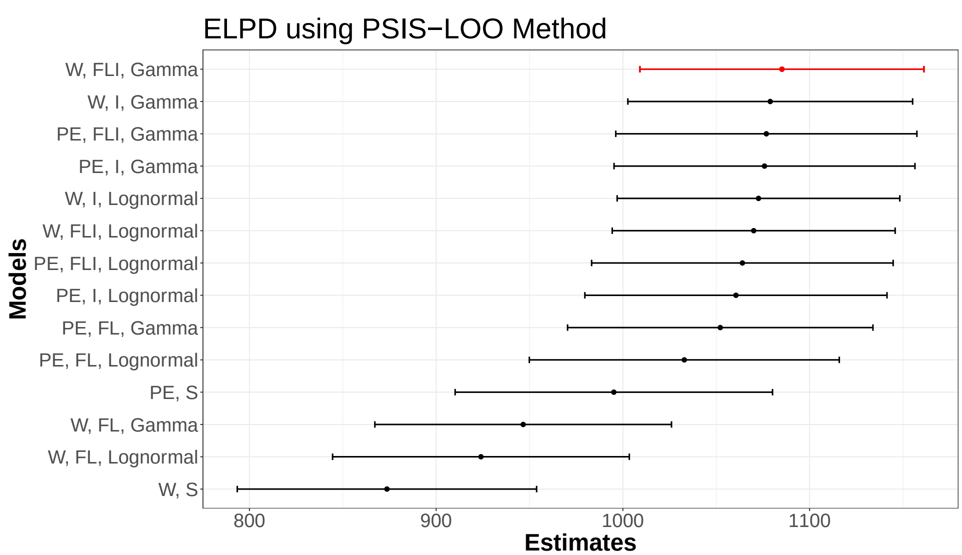


Figure B.2: Ranked comparisons of ELPD point estimates and standard errors, using the PSIS-LOO method, for human-caused wildland fires. The preferred model is highlighted in red.

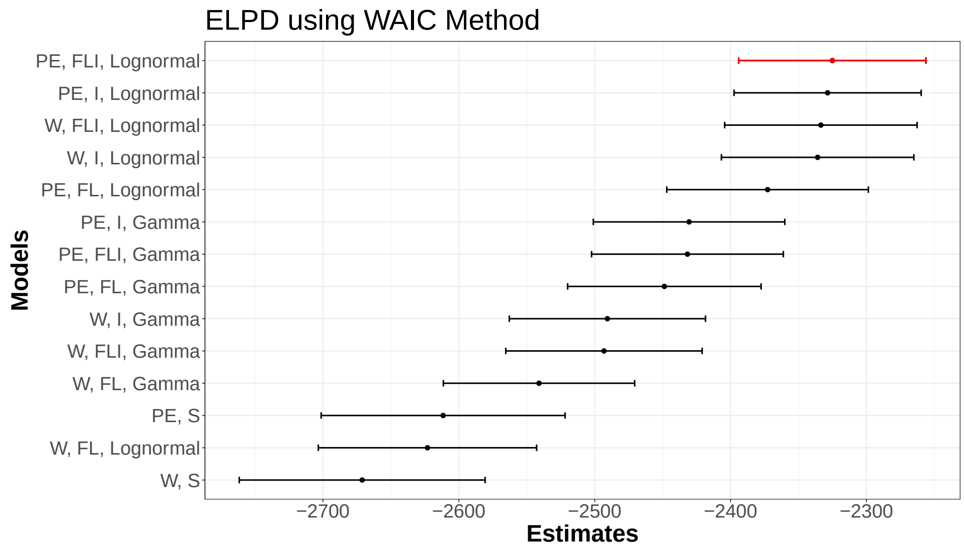


Figure B.3: Ranked comparisons of ELPD point estimates and standard errors, using the WAIC method, for lightning-caused wildland fires. The preferred model is highlighted in red.

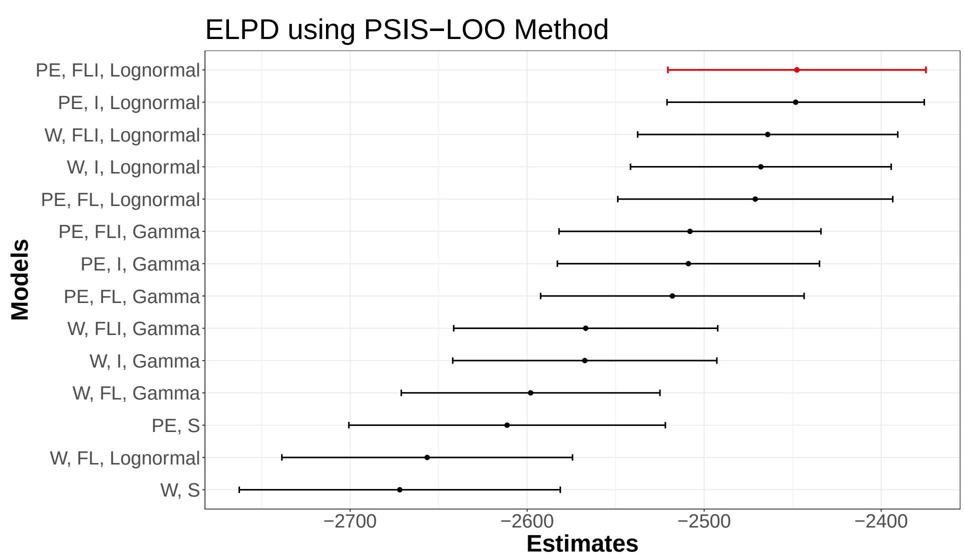


Figure B.4: Ranked comparisons of ELPD point estimates and standard errors, using the PSIS-LOO method, for lightning-caused wildland fires. The preferred model is highlighted in red.

B.3 Model Diagnostics for Lightning-Caused Fire Models

Table B.1: Summary of lightning-caused fires with $\hat{k} \geq 0.7$. The average total duration length is the sum of the average detection and action lengths, where the lengths are in days.

x	# of Fires with $\hat{k} \geq 0.7$ for x Models	Proportion of Fires with $\hat{k} \geq 0.7$ for x Models	Average Detection Length	Average Action Length	Average Total Duration Length
0	470	0.37	0.53	0.76	1.28
1	122	0.10	1.05	0.96	2.01
2	236	0.19	1.16	1.06	2.22
3	134	0.11	2.17	1.23	3.40
4	81	0.06	2.60	1.08	3.68
5	77	0.06	2.75	1.62	4.36
6	37	0.03	2.32	2.43	4.75
7	30	0.02	3.98	2.16	6.14
8	20	0.02	5.23	2.12	7.34
9	18	0.01	6.38	1.42	7.80
10	15	0.01	8.45	4.10	12.55
11	19	0.01	5.86	7.15	13.01
12	9	0.01	17.41	3.58	20.99
13	1	0	0.71	22.04	22.75
14	1	0	0.38	15.32	15.71

Table B.2: Summary of lightning-caused fires with $\hat{k} \geq 0.7$. These models were fit using 37% of the original data. The average total duration length is the sum of the average detection and action lengths, where the lengths are in days.

y	# of Fires with $\hat{k} \geq 0.7$ for y Models	Proportion of Fires with $\hat{k} \geq 0.7$ for y Models	Average Detection Length	Average Action Length	Average Total Duration Length
0	366	0.78	0.37	0.49	0.86
1	46	0.10	0.94	0.81	1.75
2	28	0.06	1.15	3.09	4.24
3	24	0.05	1.34	1.09	2.43
4	3	0.01	1.08	0.52	1.60
5	1	0	0.25	1.02	1.27
6	2	0	0.65	11.42	12.06

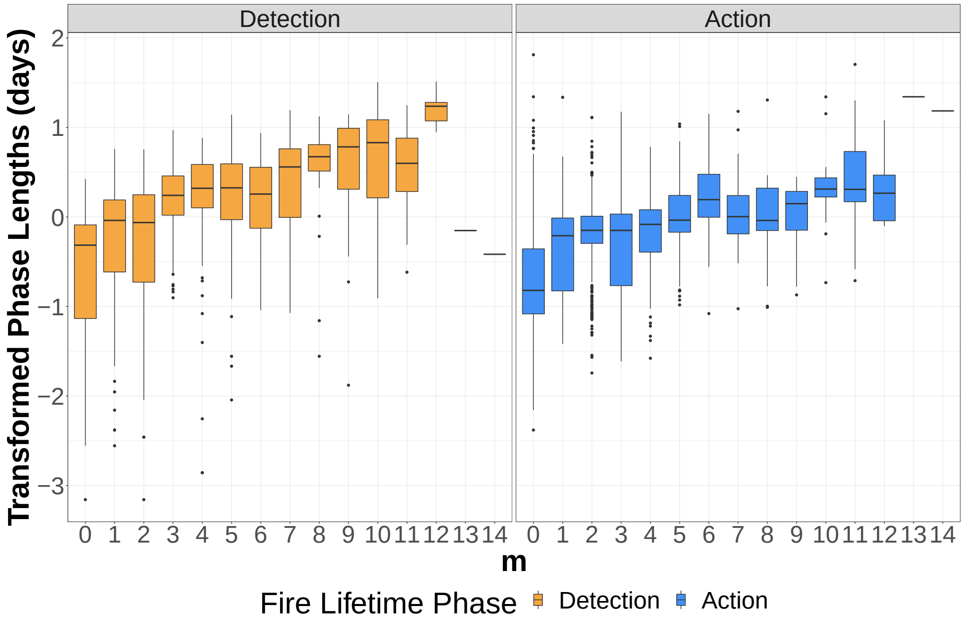


Figure B.5: Boxplot showing the spread of the log 10 transformed and stratified phase lengths (days) of the associated fires against x from Table B.1.

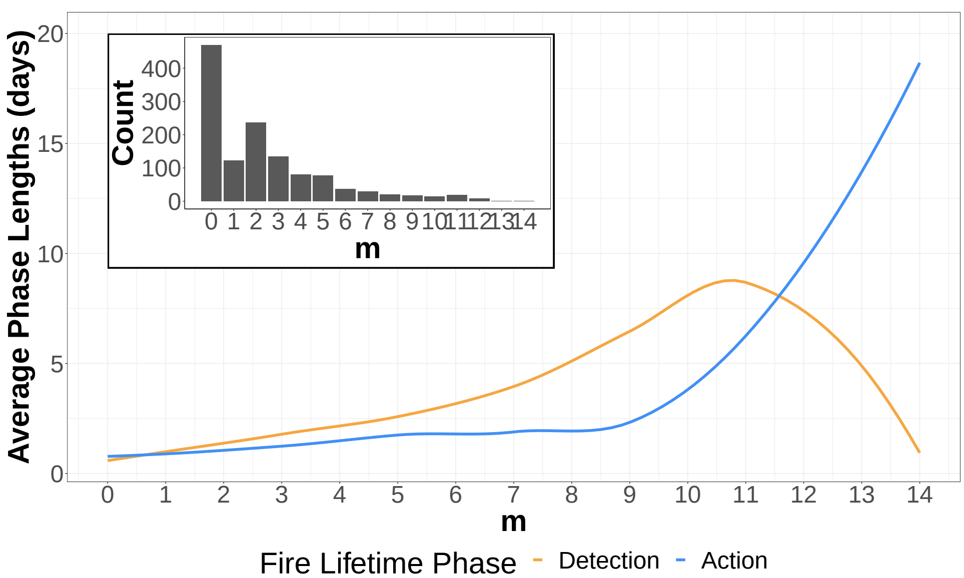


Figure B.6: Smooth lines using the locally estimated scatterplot smoothing (LOESS) method to visualize the stratified average phase lengths (days) of the associated fires against x from Table B.1. A histogram inset is provided to show the fire counts.

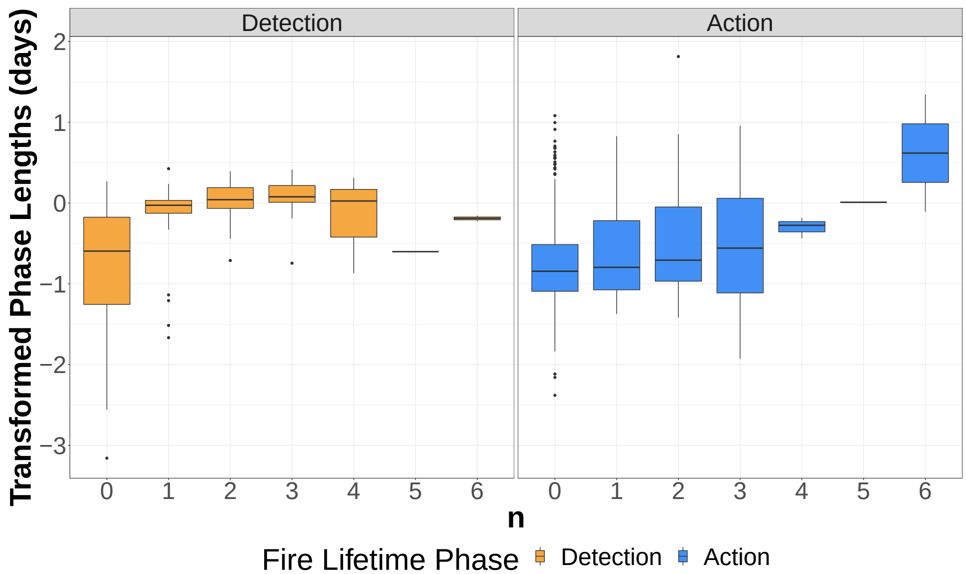


Figure B.7: Boxplot showing the spread of the log 10 transformed and stratified phase lengths (days) of the associated fires against y from Table B.2.

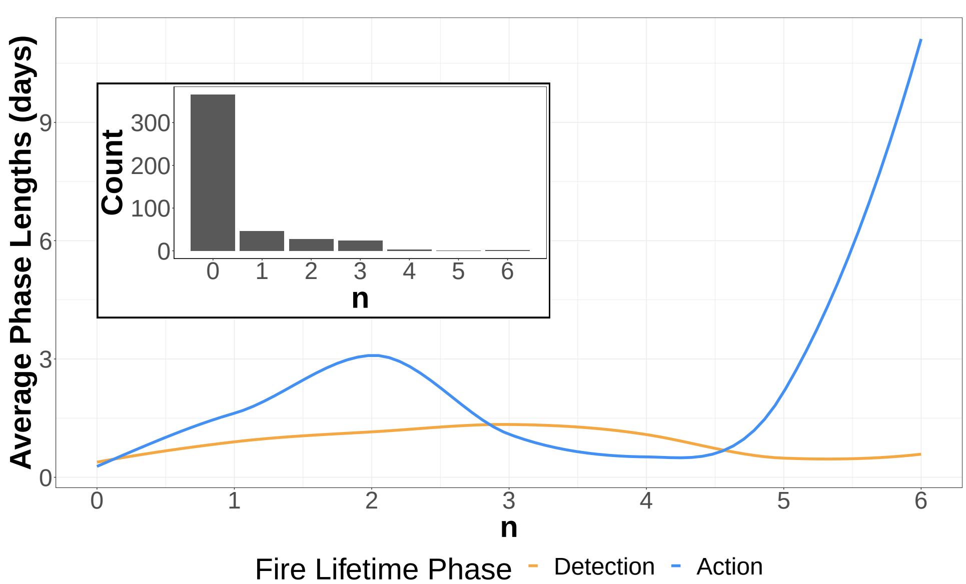


Figure B.8: Smooth lines using the locally estimated scatterplot smoothing (LOESS) method to visualize the stratified average phase lengths (days) of the associated fires against y from Table B.2. A histogram inset is provided to show the fire counts.

B.4 Simulation Study for Frailty Forms

B.4.1 FLI Form: Visualizations

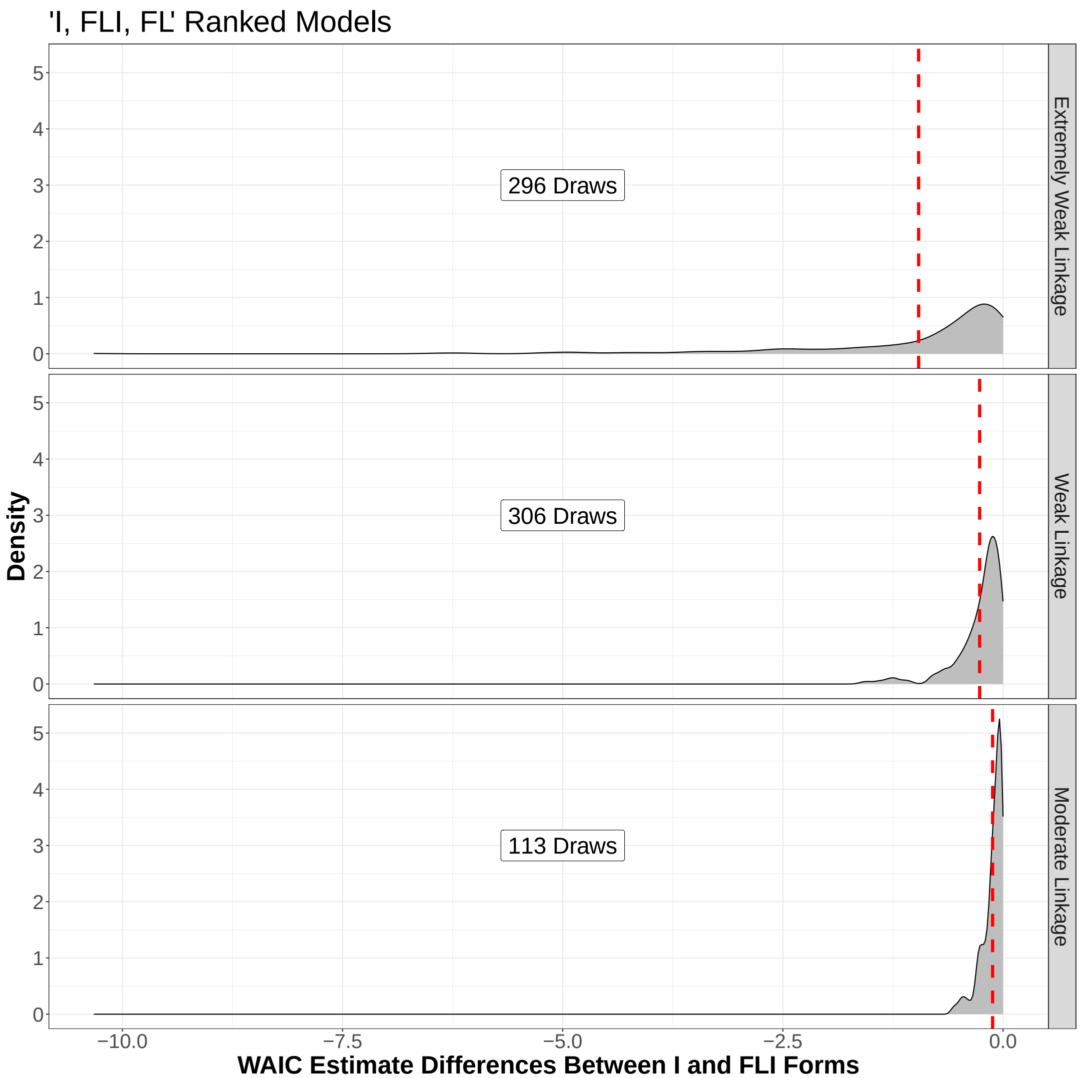


Figure B.9: Density plots of the WAIC estimate differences between the I and FLI forms using the 'I, FLI, FL' ranked models. The red dashed line represents the respective means of the differences.

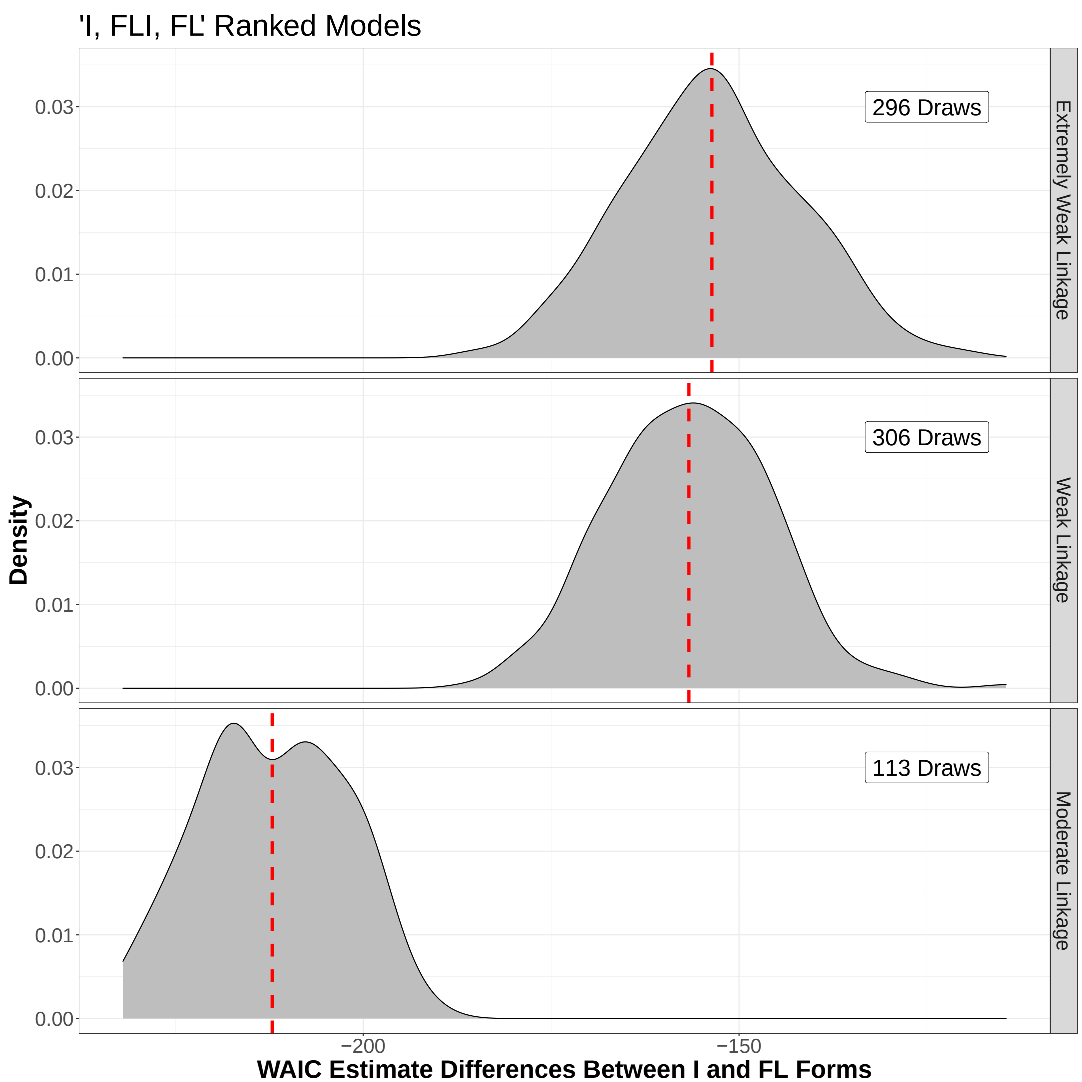


Figure B.10: Density plots of the WAIC estimate differences between the I and FL forms using the 'I, FLI, FL' ranked models. The red dashed line represents the respective means of the differences.

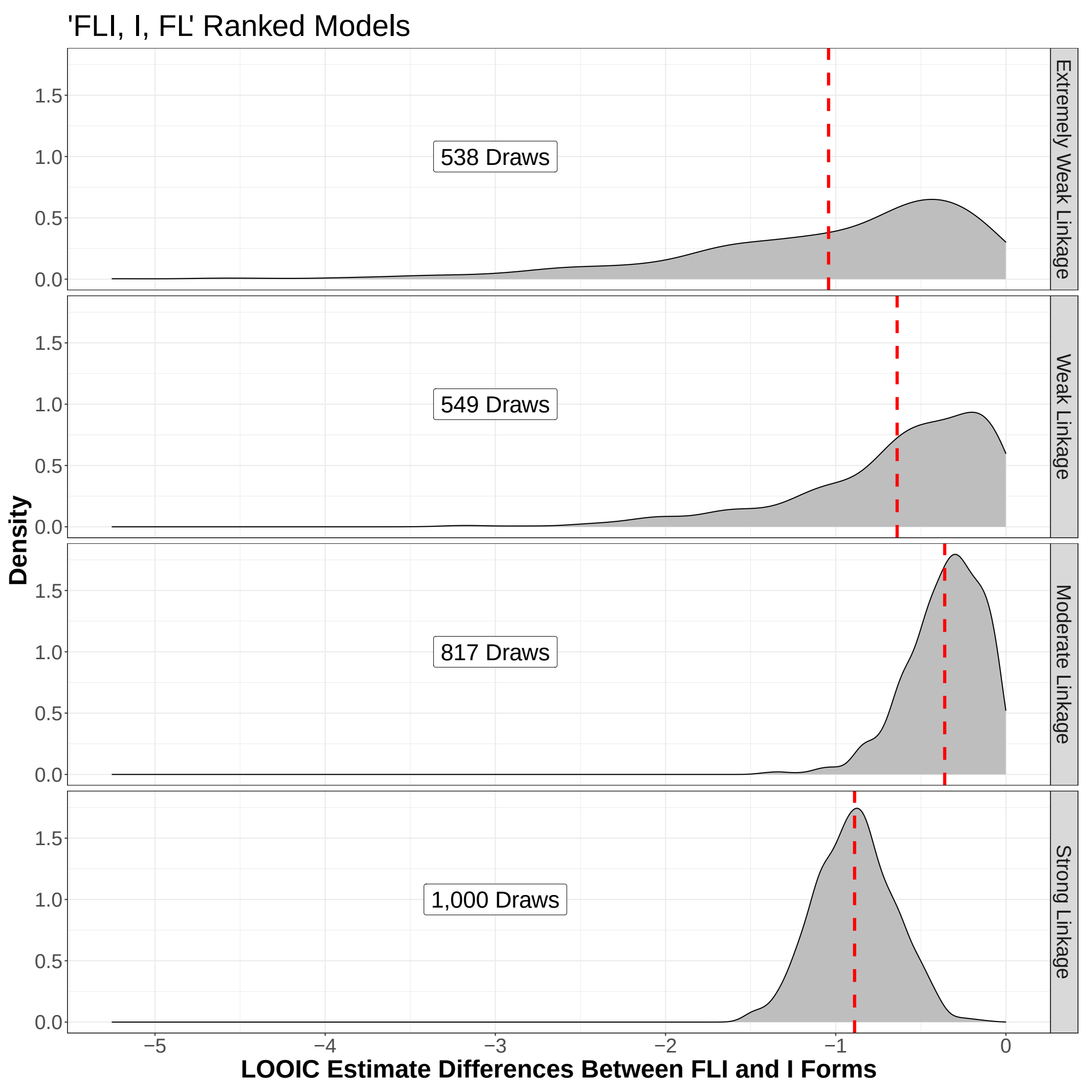


Figure B.11: Density plots of the LOOIC estimate differences between the FLI and I forms using the 'FLI, I, FL' ranked models. The red dashed line represents the respective means of the differences.

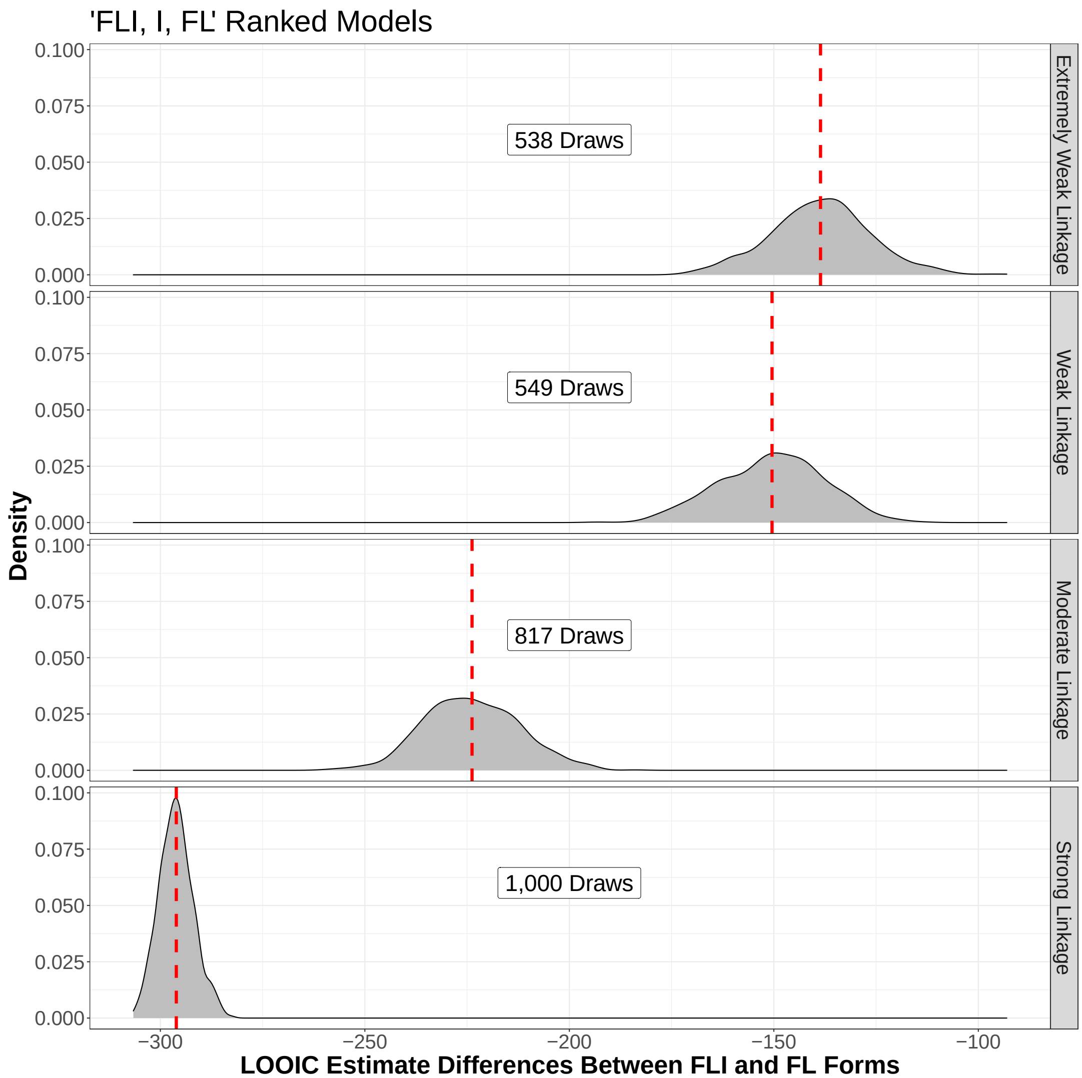


Figure B.12: Density plots of the LOOIC estimate differences between the FLI and FL forms using the 'FLI, I, FL' ranked models. The red dashed line represents the respective means of the differences.

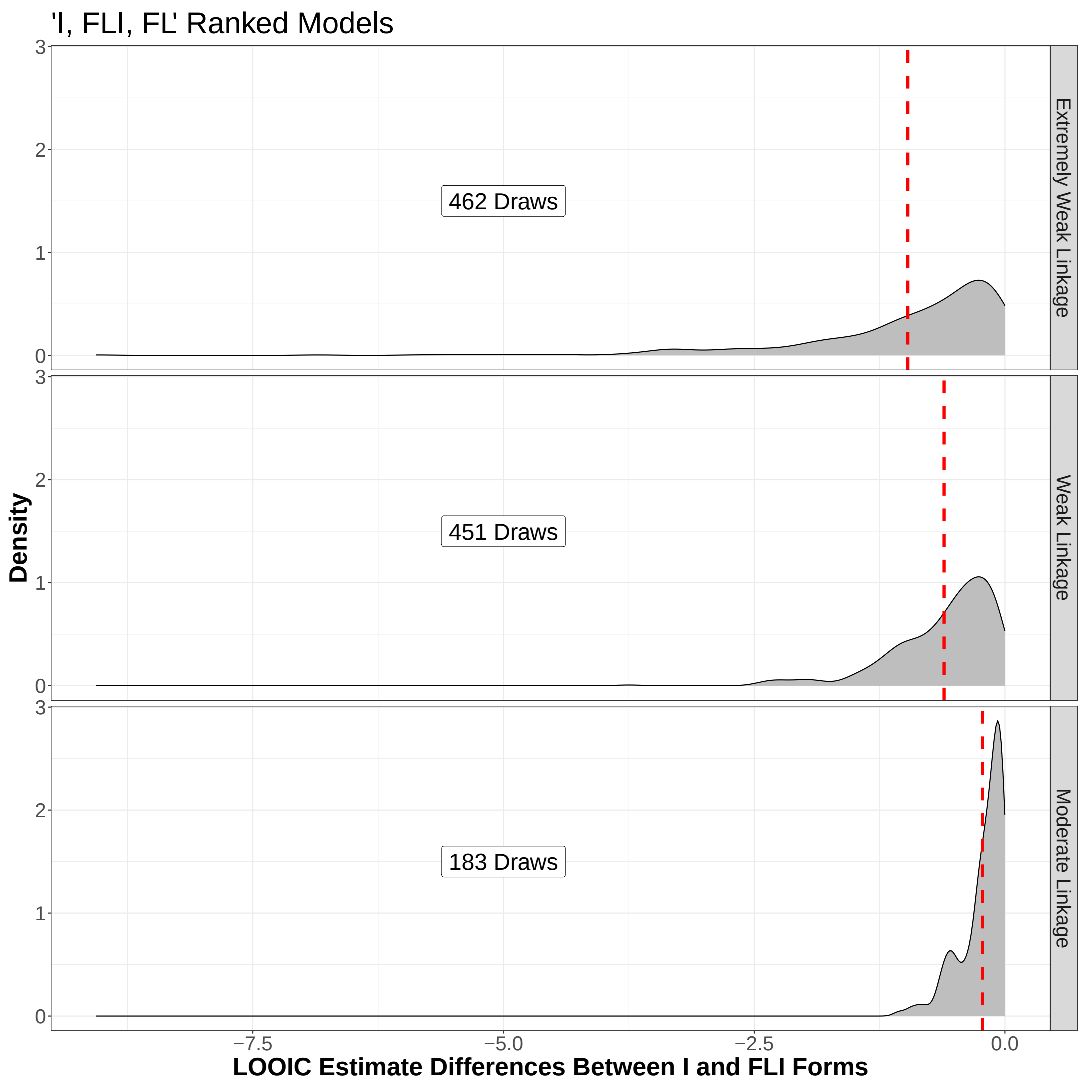


Figure B.13: Density plots of the LOOIC estimate differences between the I and FLI forms using the 'I, FLI, FL' ranked models. The red dashed line represents the respective means of the differences.

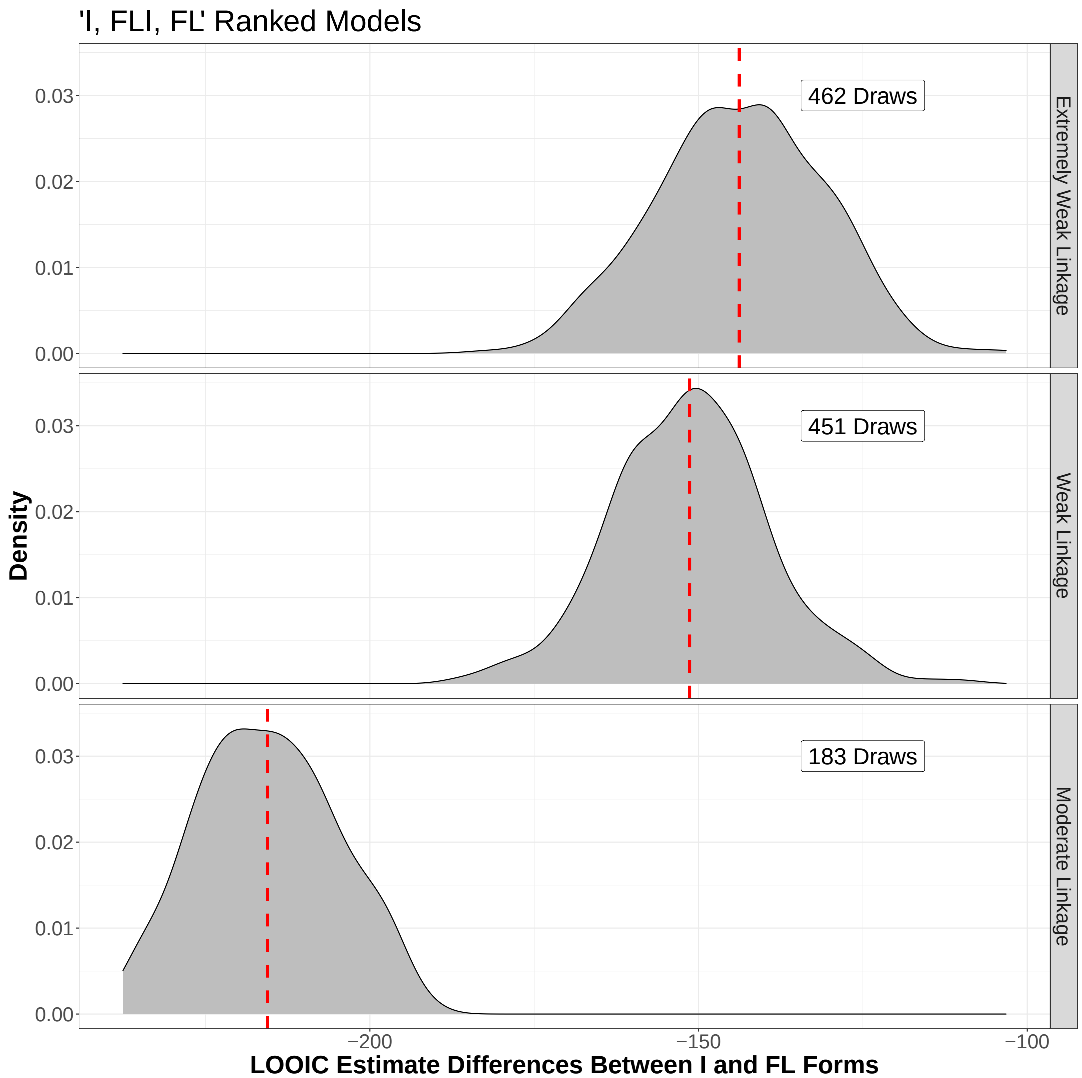


Figure B.14: Density plots of the LOOIC estimate differences between the I and FL forms using the 'I, FLI, FL' ranked models. The red dashed line represents the respective means of the differences.

B.4.2 FL Form: Results

Table B.3: A comparison of the ranked model forms, stratified by the ranking methods and the different linkages.

	Ranked Model Forms			Draws	
	Best	2nd Best	Worst	Counts	Percentages
WAIC Rank					
Extremely Weak Linkage	FLI	I	FL	1,000	100%
Moderate Linkage	FLI	I	FL	848	85%
	I	FLI	FL	152	15%
Strong Linkage	I	FLI	FL	747	75%
	FLI	I	FL	248	25%
	FLI	FL	I	5	0%
PSIS-LOO Rank					
Extremely Weak Linkage	FLI	I	FL	1,000	100%
Moderate Linkage	FLI	I	FL	764	76%
	I	FLI	FL	236	24%
Strong Linkage	I	FLI	FL	545	54%
	FLI	I	FL	430	43%
	FLI	FL	I	25	3%

Appendix C

Chapter 5 Supplementary Material

C.1 Poster

The following page contains the poster on our early wildland fire lifetime research presented at the Wildland Fire Canada 2019 conference in Ottawa in November 2019.

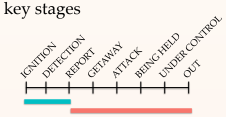
Characterizing the Lifetime Stages of Wildland Fires

Western Science

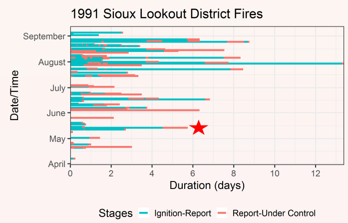
Chelsea Uggenti[1], Douglas G. Woolford[1] and Charmaine B. Dean[2]
[1]University of Western Ontario; [2]University of Waterloo

Introduction

- The lifetime of a wildland fire is made up of several key stages



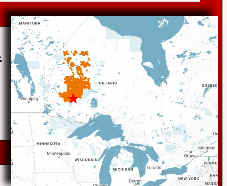
- Morin et. al. (2015) modelled the control time of fires in Ontario using a single fire lifetime stage
- We focus on two important lifetime stages:
  1. Ignition to first report
  2. First report to under control



- The duration of these stages are dependent on many variables (e.g. cause; initial attack)
- Successful initial attack fires are
  - being held within 2 days & ≤ 100ha, or
  - ≤ 4ha with no limit on time to being held

Data & Study Area

- Sioux Lookout District
- Suppressed fires only
- 1989 to 2017



Objective

Employ survival analysis methodology to determine whether the time from ignition to report has a statistically significant impact on the remaining life of a fire (up to under control)

Results

Suppression Model

For all suppressed fires, the following variables are associated with a longer report to under control stage:

- Longer **ignition-report duration**
- Higher latitude
- Higher FFMC
- Larger size at initial attack
- More fires on the landscape
- More ground forces
- More air tankers

Note: This model was stratified by three variables: cause (lightning/person), initial attack (successful/unsuccessful), and by the month of the first report date. This implies that the significant variables and effects from above were consistent across all models.

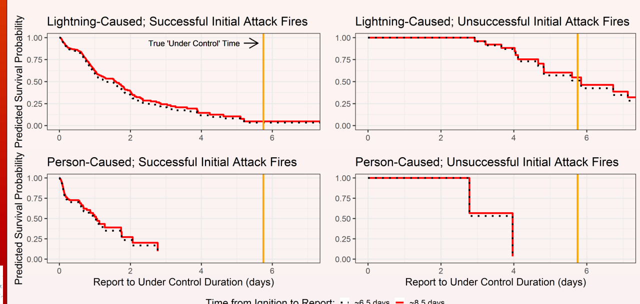
Unsuccessful Initial Attack Model

For all suppressed fires that had unsuccessful initial attack, the following variables are associated with a longer report to under control stage:

- Longer **ignition-report duration**
- Larger size at initial attack
- More fires on the landscape

*Fires that were reported later in the season lead to a shorter fire lifetime.

Note: This model was not stratified by cause (lightning/person) since there wasn't a strong difference for these types of fires.



★ Example: Two fire lifetime stages & their relationship

- The figure above shows the predicted probability curves of a fire's survival.
- Our example uses predictor settings that are based on a scenario from a fire that occurred in 1991:
  - 50.1652° latitude (see map)
  - 80 FFMC
  - 2.1ha size at initial attack
  - One other fire burning within the same district at the same time
  - 12 ground forces
  - 3 air tankers

Note: This fire was caused by lightning, had a successful initial attack, and occurred in May, 1991.

Discussion

Models

- Both models confirm a relationship between the two stages
  - Longer time spent in the ignition-report stage is associated with longer time spent in the report-under control stage
- Interesting: more ground forces and air tankers are associated with a longer fire lifetime in the suppression model
- Interesting: fires reported later in the season are associated with a shorter fire lifetime in the unsuccessful initial attack model

Example

- The shorter ignition-report stage curves (dashed black line) lie mostly to the bottom left of the longer stage curves (solid red line) in all cases
  - Confirms the relationship stated above
- Person-caused fires come 'under control' sooner than lightning-caused ones

Future Work

Extend this exploratory data analysis to study the various stages of a wildland fire and how they are related in detail using more advanced modelling frameworks

Acknowledgments

The authors gratefully acknowledge the Ontario Ministry of Natural Resources and Forestry for their data, the Natural Sciences and Engineering Research Council of Canada and Western University's Centre for Environment and Sustainability for their support, and Colin McFayden for his helpful advice.

References

- *Study area.* Snazzy Maps. (2017). Ontario, Canada. Retrieved from https://snazzymaps.com/style/118022/water-institute
- Lawless, J. F. (2011). *Statistical models and methods for lifetime data* (Vol. 362). John Wiley & Sons.
- Morin, A. A., Albert-Green, A., Woolford, D. G., & Martell, D. L. (2015). The use of survival analysis methods to model the control time of forest fires in Ontario, Canada. *International Journal of Wildland Fire, 24*(7), 964-973.

Appendix D

Chapter 6 Supplementary Material

D.1 Workshop Outline

The following pages contain the outline for the workshop developed in February-April 2021 and used for the study.

Active Learning in Math & Stats: Benefits, Limitations, and Practical Strategies for Implementing Active Learning Activities in Undergraduate University Mathematics and Statistics Courses

Graduate Teaching Assistant (GTA) Training and Development Program

Developed by:
Chelsea Uggenti, Ph.D. Candidate in Statistics, The University of Western Ontario

Summary

Many graduate teaching assistants (GTAs) begin their roles with little or no prior teaching experience. Yet they play an important role in undergraduate student learning, including the assessment processes (from marking to facilitating small discussions and so on), and are often a first point of contact for undergraduate students. As students often see GTAs as less intimidating figures than their professors, GTAs have great potential to both engage and inspire the future scholars from their disciplines (Dimitrov et al., 2013).

As noted by Gardner and Jones (2011) and references therein, both the responsibilities that science graduate teaching assistants undertake and the volume of science undergraduate courses being taught by them at research universities are increasing. Teaching training for statistics and mathematics GTAs is often limited, informal, or under-developed, and typically arises from reflection of the experience of being students themselves or "on the job" trial-and-error experiences (Gardner & Jones, 2011; Gelman, 2005). These issues are amplified when GTAs are asked to instruct introductory science, technology, engineering, and mathematics (STEM) courses since they may lack the pedagogical skills to effectively teach these courses (Crowe, 2019).

There are a variety of reasons why STEM GTAs are in need of such instructor training. Gelman (2005) notes that it may be hard for them to relate to the various types of learners in a course since graduate students are often top performers in similar environments. They are also inclined to use traditional lecture-style techniques because such approaches are familiar to them and they may have developed rigid, deeply held beliefs about teaching (Justice et al., 2017). Mathematics and statistics GTAs often resist employing active learning techniques or participatory activities in their tutorials or lectures due to anxieties that they won't have time to cover what has been identified to them as the *important* material (Gelman, 2005).

GTAs often feel overwhelmed with all the demands placed on them and their time leading to a feeling of self-preservation that is reflected in their list of priorities (Gardner & Jones, 2011). Teaching development is not placed as a top priority for most graduate students which causes them to develop their researcher identity at the expense of their instructor identity. Although knowing the content knowledge is important for effective teaching, it is problematic that neither the GTAs, nor many institutions and disciplines, prioritize pedagogical training as a requirement to take on GTA duties, including forms of instruction. Perhaps the most consistent support that

GTAs typically receive and utilize regarding their teaching comes from their peers and fellow GTA's.

The goal of this workshop is for GTAs in mathematics and statistics to gain a better understanding of active learning teaching methods. In active learning the teaching strives to involve students in the learning process more directly, rather than only utilizing the traditional lecture-style of teaching where students simply listen and take notes.

These methods are strongly needed in university mathematics and statistics courses as they promote a higher-level understanding of course concepts and often motivate students' interest (Garfield, 1993; Rosenthal, 1995). Active learning techniques have been proven to increase students' course scores and overall satisfaction (Freeman et al., 2014). Although it is necessary to continue lecture-style teaching, especially for heavily theoretical material, some form of active learning techniques can greatly benefit students' engagement of the material. Integrating active learning techniques in different environments can also be a useful way to gauge students' learning outside of a formal assessment – its an ungraded way of identifying which concepts students are struggling with and the areas that require further review.

Participants will gain necessary and vital knowledge of these methods from this workshop. By discussing tangible examples of these techniques applied to participants' home disciplines in collaboration with their peers in such disciplines, along with the benefits and limitations, it will help GTAs to start thinking about different ways to implement these ideas in their own labs, tutorials, and/or courses.

Keywords: active learning, graduate teaching assistants, mathematics, statistics

Learning Outcomes

By the end of this workshop, successful participants will be able to:

- Describe the concept of active learning, including how it differs from traditional teaching approaches.
- Explain the benefits and limitations of active learning activities used in undergraduate university mathematics and statistics courses.
- Develop strategies for implementing active learning activities in different types of environments, such as small labs/classes, large labs/classes, and online.
- Evaluate the effectiveness of active learning activities using classroom assessment techniques.

Annotated Bibliography (Chronological Order)

Garfield, J. (1993). Teaching statistics using small-group cooperative learning. *Journal of Statistics Education*, *1*(1), 1-9.

This paper provides several definitions of cooperative learning and reasons why this type of active learning is important to implement in statistics courses. Cooperative learning helps to motivate students, reinforces their understanding of the material and results in a higher level of learning the material. Garfield discusses how to implement groups, provides examples of cooperative group activities in statistics courses and mentions how to evaluate student learning. Concerns about using small groups are addressed near the end.

Angelo, T. A. (1995). Classroom assessment for critical thinking. *Teaching of psychology*, *22*(1), 6-7.

Angelo outlines that higher-ordering thinking skills like analysis, problem solving, and evaluation are difficult, but not impossible, to teach. Students can develop these skills so long as they have continuous opportunities to practice them throughout the term. Classroom assessment techniques (CATs) can be used by instructors as formative assessments to evaluate what their students know, how much they know, and how well they are learning the material. Most CATs involve student reflection or explanation of their learning. Angelo highlights the Minute Paper, performed at the end of the class, as one of the simplest CATs. Students write short, anonymous answers to the following questions: “What is the most important thing you learned in today’s class? What is one question you have from today’s class?”

Keeler, C. M., & Steinhorst, R. K. (1995). Using small groups to promote active learning in the introductory statistics course: A report from the field. *Journal of Statistics Education*, *3*(2), 1-9.

This paper offers several reasons for implementing cooperative learning techniques in undergraduate statistics courses. They describe the entire set-up of their study and provide changes that were made during the second classroom trial. Their results suggest that cooperative learning increases student’s success rate of passing the course, increases their course marks and generally increases their satisfaction in the course. Most students were more engaged in the material but still wanted some traditional lectures to assist with their understanding of the conceptual material.

Rosenthal, J. S. (1995). Active learning strategies in advanced mathematics classes. *Studies in Higher Education*, *20*(2), 223-228.

Rosenthal suggests several alternative teaching strategies that can be used to augment the typical lecture format of teaching within university mathematics courses. He begins by noting that small group exercises where students are assigned problem-solving exercises allow students to learn from each other and learn by teaching each other. Written assignments are also offered to encourage students to think more deeply about the content and to see the material in a larger context. Having these written assignments reviewed by at least two peers within the course also strengthens their understanding of the content along with written communication skills. However, less formal writing assignments can also be utilized in mathematics courses by implementing approaches like the Minute Paper at the end of a lecture.

Prins, S. C. B. (2009). Student-centred instruction in a theoretical statistics course. *Journal of Statistics Education*, *17*(3), 1-12.

Prins provides a practical example of how student-centred instruction can be utilized in a theoretical undergraduate statistics course. Students were randomly assigned one problem at the end of their teacher-centred lesson on Tuesday that they would have to present in class during the following Thursday's class time. The questions assigned were standard rather than unusual applications of a concept – the unusual application questions were saved for take-home work. The instructor wrote down the solutions that students presented, annotated any possible errors, and uploaded the correct and complete solutions on the course website. Students found that this approach helped to create a supportive community and that it reinforced their learning of course concepts.

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance is science, engineering, and mathematics. *Proceedings of the National Academy of Sciences*, *111*(23), 8410-8415.

In the STEM classroom, should we ask or should we tell? Freeman et al. (2014) sought to answer this question by performing a quantitative data analysis to determine how constructivist (i.e., learner-centred) versus exposition-centred (i.e., instructor-centred) methods impact student performance in undergraduate STEM courses. They meta-analyzed 225 studies in the published and unpublished literature, that reported data on examination scores or failure rates when comparing student performance in these courses, under traditional lecturing versus active learning. The results showed that incorporating active learning in such courses increases student examination performance and that exclusively lecturing increases student failure rates by 55%. This suggests that the student-centred approach to teaching may lead to increases in student performance which strengthens the call to include more active learning in undergraduate STEM courses.

Lang, J. M. (2016). *Small teaching: Everyday lessons from the science of learning*. John Wiley & Sons. San Francisco, USA.

This book outlines small, yet powerful, changes that can be implemented in a single class for improving student learning. Each chapter presents a concept in cognitive theory, explains when and how it should be employed, and provides examples of how the intervention can be utilized in the classroom. It also offers short tips and tricks associated with each concept. For example, Lang notes that retrieving material from memory works well when using brief classroom interventions like setting weekly, low-stakes quizzes that help students cement their foundational course content or by ending the class with a short quiz on the material learnt from that day.

Braun, B., Bremser, P., Duval, A. M., Lockwood, E., & White, D. (2017). What does active learning mean for mathematicians? *Notices of the American Mathematical Society, 62*(2), 124-129.

This notice defines active learning and provides examples of active learning techniques and environments. They discuss classic examples like Think-Pair-Share, classroom response systems (or iClickers), and flipped classes, along with discipline-specific examples like inquiry-based learning (IBL). IBL is a common active learning technique in mathematics since class time is spent with students working individually or in groups on problems and providing solutions or proofs to the class. They outline several things to expect when incorporating active learning techniques and addresses many concerns that mathematics instructors may have when thinking about utilizing active learning in their courses.

Kerrigan, J. (2018). Active learning strategies for the mathematics classroom. *College Teaching*, *66*(1), 35-36.

This paper provides three "quick fixes" for making undergraduate mathematics courses more active and engaging. It notes that one way of creating a welcoming learning environment is to develop the course syllabus in conjunction with the students; that way, students have an opportunity to discuss their goals, concerns, and any supports they will need and instructor's have the chance to share their expectations for the students, assessment methods, and so on. Kerrigan suggests using gallery walks or class polling to elicit prior knowledge as part of a pre-assessment review of concepts. It is also suggested that online games (like Kahoot! or Quizizz) is an efficient method for formative assessment and reviewing material prior to a midterm or exam.

Content and Organizational Table

Note: All active learning activities used within this workshop are in *italics*.

Duration (mins)	Subject	Activity	Purpose
5	Motivating Activity (*Dotmocracy*)	Part A: Write statements about various thoughts on active learning (see Appendix A) on chart paper and place them around the room. Provide participants with a set number of dot stickers and ask them to walk around the room, thinking about each of the statements and putting one or more dots on the statements they most strongly agree with (more notes = more significance). Part B: Discuss the results of the activity with the entire group by visually assessing the number of dot stickers under each statement.	Motivate participants' interest in the workshop and provide a "bridge-in" for the rest of the workshop content. Gauge participants' initial beliefs on active learning in mathematics and statistics courses. Acts as an "icebreaker" activity by strengthening bonds between participants (some may be new graduate students) and reducing any tension.
35	Introduction & Learning Outcomes (*Quescussion; Think-Pair-Share*)	Provide a brief overview of the need for active learning activities in undergraduate university mathematics and statistics courses. Present the learning outcomes for the workshop along with the definition, benefits, and limitations of active learning activities in a PowerPoint presentation. Also, present methods for evaluating active learning activities using Classroom Assessment Techniques (CATs). The presentation will provide concrete examples of active learning activities that GTAs can incorporate into tutorials, labs, office hours, or help desk hours for math and stats courses (Quescussion for extrapolation of runners' times example; Think-Pair-Share for incorrect proof examples).	Share the learning outcomes to clarify what participants will get out of the workshop. Explain the foundations of active learning activities (with cited examples). Stimulate interest in these techniques by focusing on the benefits of active learning activities for the students and the instructors/GTAs. Examine the limitations that arise when implementing active learning activities in math and stats courses and discuss scenarios where these limitations may be mitigated or must be accepted.

			Encourage participants to put themselves in students' shoes and explore how active learning activities can be used effectively to explain difficult concepts in math and stats by having them participate in two activities.
52	Participatory Learning (*Café Conversation*)	Participants will be placed into three, six, or nine groups depending on numbers. Each group will represent a different learning environment that GTA's may encounter (e.g., lab/tutorial, office hour/help desk, and an online class) for a fictional undergraduate math or stats course. Part A: Give participants a few minutes to reflect on any active learning techniques they have previously seen or experienced in courses or even in this workshop. (Time: 1-2 mins) Part B: Each group writes down any examples of active learning activities, and brief explanations of them, that they think would be helpful to incorporate in their specific learning environment (i.e., lab/tutorial, office hour/help desk, or an online class). (Time: 10 mins) *Groups are asked to move tables clockwise. (Time: 30 secs) Part C: Each group writes down any benefits and limitations for the active learning examples stated within that learning environment. (Time: 10 mins) *Groups are asked to move tables clockwise. (Time: 30 secs)	Use small groups to foster collaboration and community in the workshop. Assess participants' prior knowledge and experience of active learning activities within undergraduate math and stats courses. Entice participants to (partially) set aside the traditional practice of lecturing and to develop concrete active learning activities that could be utilized in a variety of learning environments. Critique the active learning activities that their peers suggested in a constructive manner to develop approaches for overcoming any limitations. Distinguish between activities that are believed to be helpful when learning math/stats and those that are not. Analyze any differences or patterns between these activities and discuss their shortcomings in the learning process. Assess participants' learning and provide constructive feedback on their suggestions.

		Part D: Each group writes down ideas of how they could possibly overcome those limitations and how they will evaluate the active learning activities within their learning environment. (Time: 10 mins) *Groups are asked to move tables clockwise. (Time: 30 secs) Part E: Have each group return to their initial learning environment to read and discuss all the comments and ideas that their peers provided. (Time: 5 mins) Part F: Debrief the activity with the entire group. Have each group nominate one person to give a summary of the examples, benefits, limitations, strategies for overcoming limitations, and assessment methods for implementing active learning activities in their specific learning environment. (Time: 15 mins)	Debrief the activity as an entire group to highlight the range of possible active learning activities that can be applied to various environments.
5	Post-Assessment (*1-Minute Paper*)	Provide participants with two questions for brief reflection: 1. What was the most useful or meaningful thing you learned from this workshop? 2. How do you think you will use what we learned today moving forward? Emphasize that responses should be concise. Have each participant record and anonymously submit their answers.	Assess participants' learning and understanding of the workshop contents. Determine if there are any gaps in their understanding that needs to be addressed.
5	Summary	Distribute final handout (see Appendix E). Summarize the main points of the workshop and any "take away" ideas. Meta-moment: Draw attention to the various active learning activities used throughout this workshop.	Direct participants towards available resources. Encourage them to incorporate the ideas presented in this workshop into their current or future TA-ships and teaching.
Total Time: 1 hour, 42 minutes			

Presentation Strategies

Teaching development workshops have several benefits for GTAs and for the undergraduate students that they interact with. Fostering the growth of the teaching ability of GTAs can also enhance their abilities to communicate their research and help them become better researchers in general since many of the communication and other skills they develop are transferrable. Therefore, training programs on effective teaching and pedagogy add to the overall professional development of graduate students and should boost their credibility as a researcher.

Teaching workshops can also serve as a gateway to further teaching development since they often help participants understand the benefits that can be gained from them. For instance, workshops geared towards new GTAs frequently introduce skills that may be strengthened if they continue in teaching development throughout their careers. It provides them with an opportunity for career development and can be included in their CV as a professional activity.

Throughout the course of the workshop, it is important to draw attention to the many "meta-moments" of active learning techniques. The facilitator should pause after each activity and remark on what the facilitator and participants are doing and why. For example, the facilitator can mention that the dotmocracy activity requires some preparation (e.g., questions to ask, supplies required, etc.) and little time versus the Think-Pair-Share activity that requires almost no preparation and minimal time. During the meta-moments, the facilitator also needs to discuss how one might adapt certain activities to work online. For example, a Think-Pair-Share activity in an online environment could be done by sending students into breakout rooms for the "Pair" component.

This workshop is designed to take place in a classroom equipped with a data projector. Ideally, participants should be seated at tables around the room to facilitate group work for the Café Conversation activity. We suggest holding this workshop in an interactive classroom like the Western Active Learning Spaces (WALS); see https://www.uwo.ca/wals/ for more details. Materials needed for this workshop include chart paper, coloured markers, coloured pens, dot stickers, cue cards, and paper handouts.

Activity 1: Motivating interest (Dotmocracy)

This activity acts as a bridge in for the rest of the workshop. It allows the instructor to build motivation for the workshop content, gain the attention of the participants, and helps to establish the relevance of the workshop. As discussed above, there are many common concerns about active learning activities including the historical dominance of the lecture format, being unable to cover the same amount of material, and whether it is even necessary (many mathematicians and statisticians have not personally experienced undergraduate teaching environments that include active learning components).

Make sure to hand out dot stickers to participants (4/person) in advance of the activity. The statements with the most dots "win" and have more significance for this group of individuals. During the debrief of the activity, be sure to respond to the outcome of each statement and respond with "if, then" comments. For example, suppose that the statement "I've never tried active learning before and I'm very skeptical" has the most dots under it. The workshop facilitator may respond to this result by saying: "If you have never tried active learning before and skeptical about the whole thing, then you are certainly not alone! Clearly, many people feel the same way as you do. The purpose of this workshop is to introduce you to different active learning activities that you can easily add to your courses or labs/tutorials and they are all backed up by research!".

Be warned that the Dotmocracy activity may be broken. This type of voting is not reliable because of the bandwagon effect and how easy it is to mess up or cheat. Although it is quick and fun, don't fully trust the results you encounter.

Activities 2 & 3: Introduction & Learning Outcomes (Think-Pair-Share; Quescussion)

In this section of the workshop we present the three learning outcomes and introduce the definition, benefits, and limitations of active learning activities. We will also present methods for evaluating active learning, mainly using Classroom Assessment Techniques (CATs). Examples of simple, common CATs will be provided. As these ideas may be new to participants, it is important to give clear definitions and examples.

The presentation will end with two concrete examples of active learning activities that can used by GTAs in undergraduate mathematics or statistics courses.

<u>Quescussion Example</u>

A figure of the world record times in the mile run for men from 1900 to 1950 will be shown (see Appendix B1, Figure 1). Explain that you will discuss how well a straight line fits the data through questions only (i.e., participants may only respond or add to the discussion in the form of more questions). Share the three rules of the activity:

1. Only questions are allowed.
2. If someone makes a statement everyone yells "statement!"
3. Two other people must speak before a participant can speak again.

The facilitator starts the Quescussion by asking the questions: "Does fitting the data with a straight line give a reasonable prediction for the year 2100? What about the year 2000?".

Afterward, reveal the right half of the curve (see Figures 2 and 3 in Appendix B1). Discuss how the linear extrapolation works relatively well all the way to the year 2000, but that you wouldn't expect this to work out to the year 2050 and beyond.

Debrief the activity by noting that further discussion can focus on one or two of the key questions raised in greater depth. Also relay to participants that this activity works well with dense or difficult material since one has to ask genuine questions about what is going on. It also provides a less intimidating, low-stakes environment where people can make mistakes since you are asking them to generate a variety of thoughts about the topic without having them directly state their views or solve the problem.

Think-Pair-Share Example

Hand out papers containing incorrect proofs/examples (see Appendix B2) to participants; each participant should receive one paper. Ask participants to independently annotate on the paper where there are errors in the proof or example using a different coloured pen. Then have each participant pair up with someone around them to discuss their ideas or answers for each of their handouts. Bring everyone together to discuss the answers.

Debrief the activity by noting that annotating students' work in this way (during office hours or help desk hours, for example) can show them, in real time, how a correct or incorrect solution would still lose marks due to a lack of completeness.

Activity 4: Participatory Learning (Café Conversation)

During this section of the workshops, participants have an opportunity to be actively involved in achieving all three of the learning outcomes. It will help them improve their understanding of active learning and deepen their knowledge of which activities are applicable for various learning environments.

The Café Conversation is a cooperative group activity in which participants are interdependent to achieve a common goal. In each phase of the activity, participants are provided with a prompt and asked to assess it for a different learning environment. By the end, each person will have had an opportunity to think about active learning in a different environment. The success of every group depends on each individual and therefore emphasizes engagement from individual participants.

It is important that the facilitator explains the concept of the café activity at the start and how participants will move around the room (see Appendix C). Also give participants 1-2 minutes before the activity to reflect on any active learning techniques they have seen or experienced in their previous undergraduate or graduate courses. You may pose the following questions:

- What techniques did you like as a student?
- How did those techniques effect your learning or understanding of the course material?
- Were there any techniques that you did not like? Why?

During the activity the facilitator will be required to keep track of time and give participants warnings of when that round is ending (e.g., one-minute left warning). The facilitator is also expected to walk around the room observing and listening to the different groups – they may pose leading questions or insightful statements to help any groups that are finding the task difficult. An example of a completed chart for a tutorial or lab is provided in Appendix D.

Activity 7: Post-Assessment (1-Minute Paper)

This activity gives the facilitator an opportunity to find out what the participants have learned and if there are any gaps in their understanding of the workshop material. Ending the workshop with this short writing activity is a powerful way to assess the degree to which the participants understood the presented material and provides them with an opportunity to reflect on how they will incorporate the material in their future courses or labs/tutorials.

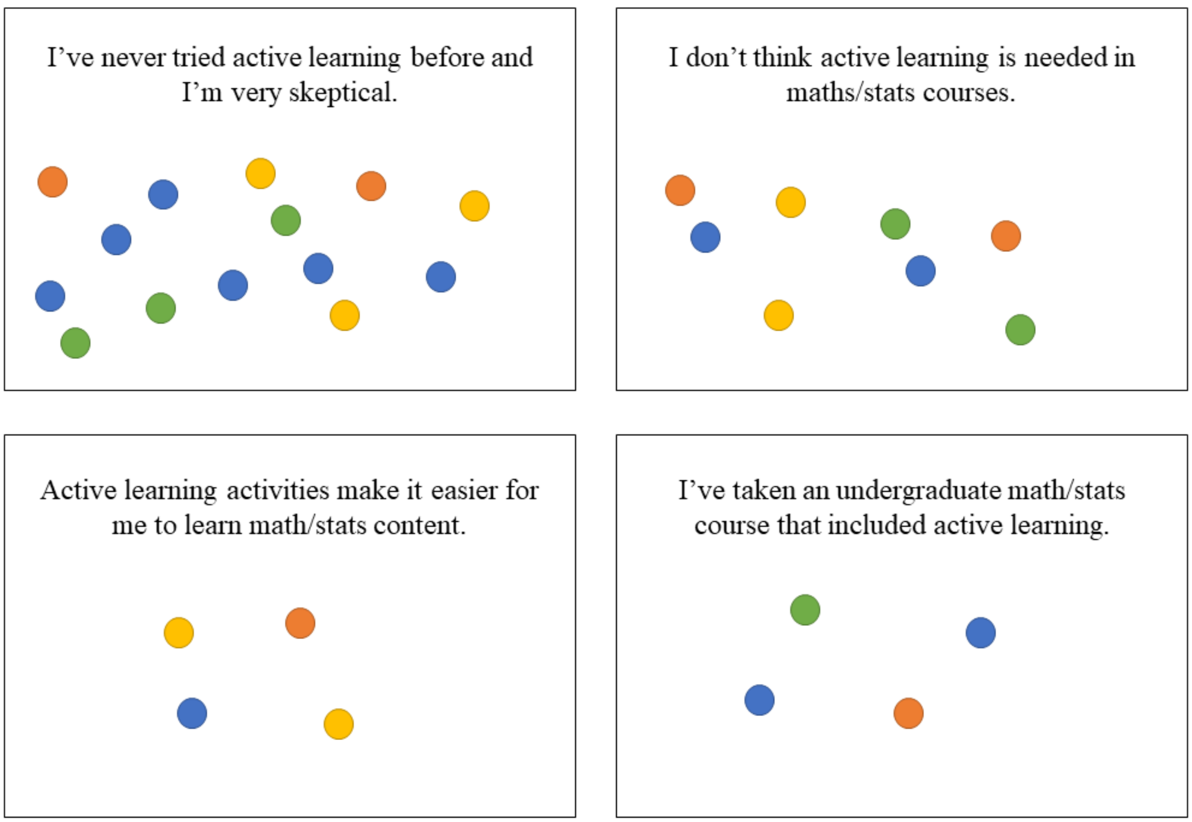
Appendix A

Dotmocracy Activity

The following statements on active learning will be used for the activity:

- "I've never tried active learning before and I'm very skeptical."
- "I don't think active learning is needed in math/stats courses."
- "Active learning activities make it easier for me to learn math/stats content."
- "I've taken an undergraduate math/stats course that included active learning."

The four pieces of chart paper with each statement will be placed on the walls around the room. An example of how the chart papers will be filled out by the end of the activity is shown below.



Appendix B1

Quescussion Activity

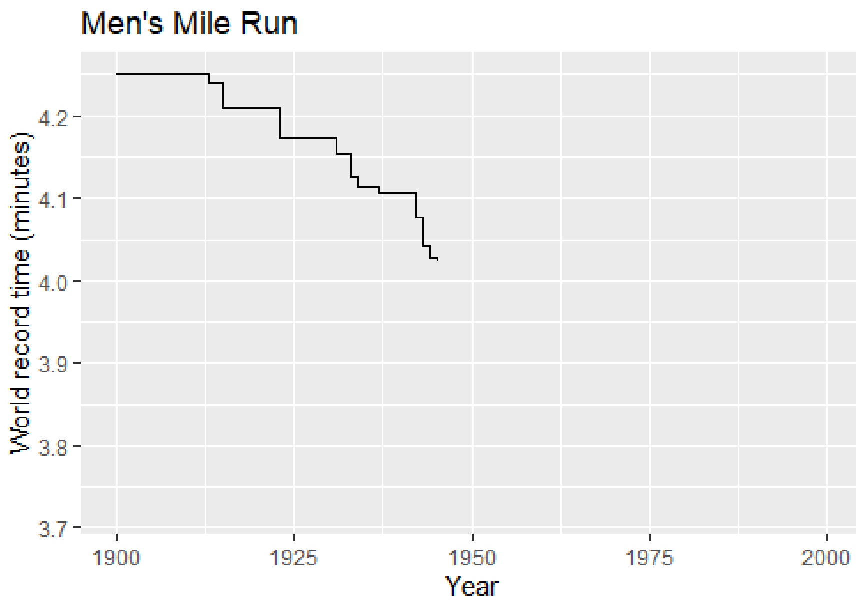


Figure 1: First plot to show to participants. Get them to discuss whether a straight “line-of-best-fit” works for year 2100 (or year 2000).

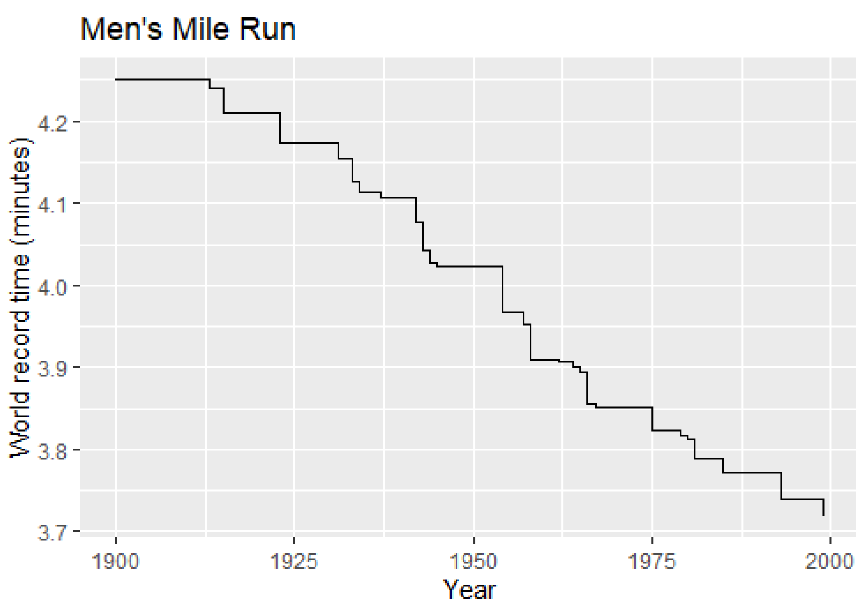


Figure 2: Second plot to show to participants of the entire dataset from 1900-2000.

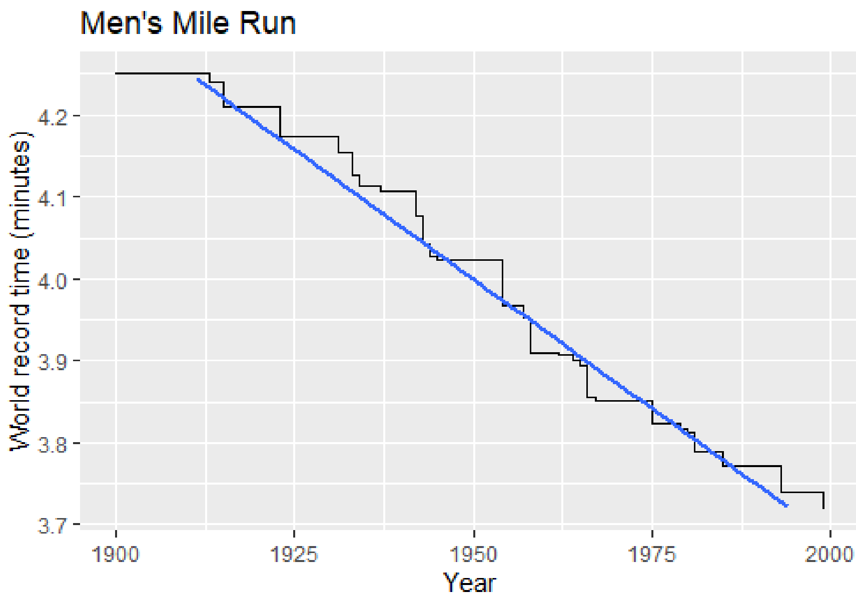


Figure 3: Final plot to show to participants. How does this “line-of-best-fit” compare to what was discussed during the activity?

Appendix B2

Think-Pair-Share Activity

Pages 17-20 show the four examples that will be used in this activity. The proofs used in the examples were taken from the following courses:

1. Calculus I
2. Calculus II
3. Introduction to Probability
4. Regression

Example 1:

Calculus I Example

If $4x-9 \leq f(x) \leq x^2-4x+7$, for all $x \geq 0$, find $\lim\limits_{x\to 4} f(x)$.

Sol^n:

$$\begin{aligned} &\lim 4x-9 \\ &= 4(4)-9 \\ &= 16-9 \\ &= 7 \end{aligned} \qquad\qquad \begin{aligned} &\lim x^2-4x+7 \\ &= (4)^2-4(4)+7 \\ &= 7 \end{aligned}$$

$$\therefore \lim f(x) = 7$$

Sol^n with corrections:

$$\begin{aligned} &\lim_{x\to 4} 4x-9 \\ &= 4(4)-9 \\ &= 16-9 \\ &= 7 \end{aligned} \qquad\qquad \begin{aligned} &\lim_{x\to 4} x^2-4x+7 \\ &= (4)^2-4(4)+7 \\ &= 7 \end{aligned}$$

$$\because \lim_{x\to 4} 4x-9 = \lim_{x\to 4} x^2-4x+7 = 7$$

$$\therefore \lim_{x\to 4} f(x) = 7$$, by the Squeeze Theorem

Example 2:

Calculus II Example

Find the length of the polar curve $r = \theta^2$ for $0 \leq \theta \leq 2\pi$.

Sol$^{\underline{n}}$:

$$
\begin{aligned}
L &= \int_\alpha^\beta \sqrt{[f(\theta)]^2 + [f'(\theta)]^2}\, d\theta \qquad , f(\theta) = \theta^2 \\
&= \int_0^{2\pi} \sqrt{(\theta^2)^2 + (2\theta)^2}\, d\theta \\
&= \int_0^{2\pi} \sqrt{\theta^4 + 2\theta^2}\, d\theta \\
&= \int_0^{2\pi} \sqrt{\theta^2(\theta^2 + 2)}\, d\theta \\
&= \int_0^{2\pi} \theta \sqrt{\theta^2 + 2}\, d\theta \qquad u = \theta^2 + 2 \\
&= \tfrac{1}{2} \int_2^{4\pi^2+2} \sqrt{u}\, du \\
&= \tfrac{1}{2} \cdot \frac{u^{3/2}}{3/2} \Bigg|_2^{4\pi^2+2} = \tfrac{1}{3}(4\pi^2+2)^{3/2} - \tfrac{1}{3}(2)^{3/2} \\
&\approx 88.10
\end{aligned}
$$

Sol$^{\underline{n}}$ with corrections:

$$
\begin{aligned}
L &= \int_\alpha^\beta \sqrt{[f(\theta)]^2 + [f'(\theta)]^2}\, d\theta \qquad , f(\theta) = \theta^2 \text{ and } f'(\theta) = 2\theta \\
&= \int_0^{2\pi} \sqrt{(\theta^2)^2 + (2\theta)^2}\, d\theta \\
&= \int_0^{2\pi} \sqrt{\theta^4 + \cancel{2}\,4\theta^2}\, d\theta \\
&= \int_0^{2\pi} \sqrt{\theta^2(\theta^2 + 2)}\, d\theta \\
&= \int_0^{2\pi} \theta \sqrt{\theta^2 + 2}\, d\theta \qquad \text{, as } \sqrt{\theta^2} = \theta \text{ since } \theta \geq 0 \qquad u = \theta^2 + \cancel{2}\,4 \\
&= \tfrac{1}{2} \int_2^{4\pi^2+2} \sqrt{u}\, du \qquad du = ? \\
&= \tfrac{1}{2} \cdot \frac{u^{3/2}}{3/2} \Bigg|_2^{4\pi^2+2} = \tfrac{1}{3}(4\pi^2+2)^{3/2} - \tfrac{1}{3}(2)^{3/2} \\
&\approx 88.10
\end{aligned}
$$

$\therefore$ statement?

Example 3:

Introduction to Probability Example

Suppose A and B are independent events and that $P(B) > 0$. Prove that A′ and B are independent events.

Sol^{n}:

$$B = (A \cap B) \cup (B \cap A')$$
$$P(B) = P(A \cap B) + P(B \cap A')$$
$$\Rightarrow P(A' \cap B) = P(B) - P(A \cap B)$$
$$= P(B) - P(A)P(B)$$
$$= P(B)P(A')$$

$\therefore$ A′ and B are independent

Sol^{n} with corrections:

$$B = (A \cap B) \cup (B \cap A')$$
$$P(B) = P(A \cap B) + P(B \cap A') \text{, since they are mutually exclusive}$$
$$\Rightarrow P(A' \cap B) = P(B) - P(A \cap B)$$
$$= P(B) - P(A)P(B), \quad \because P(A \cap B) = P(A)P(B)$$
$$= P(B)P(A'), \quad \because 1 - P(A) = P(A')$$

$\therefore$ A′ and B are indepent

Example 4:

Regression Example

Consider the simple linear regression model $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$, $i = 1, \dots, n$ where the intercept β_0 is known and $E(\varepsilon_i) = 0$, $Var(\varepsilon_i) = \sigma^2$, $i = 1, \dots, n$. Find the variance of $\hat{\beta}_1 = \dfrac{\sum_{i=1}^n (y_i - \beta_0) x_i}{\sum_{i=1}^n x_i^2}$.

Sol$^{\underline{n}}$:

$$
\begin{aligned}
Var(\hat{\beta}_1) &= Var[\Sigma (y_i - \beta_0) c_i] \\
&= \Sigma c_i^2 \, Var(y_i - \beta_0) + \Sigma\Sigma \, Cov[(y_i - \beta_0), (y_j - \beta_0)] \\
&= \Sigma c_i^2 \, Var(y_i - \beta_0) \\
&= \Sigma c_i^2 \, Var(y_i) \\
&= \sigma^2 \, \Sigma \left(\frac{x_i}{\Sigma x_i^2} \right)^2 \\
&= \sigma^2 \, \frac{\Sigma x_i^2}{\Sigma x_i^4} \\
&= \frac{\sigma^2}{\Sigma x_i^2}
\end{aligned}
$$

Sol$^{\underline{n}}$ with corrections:

$c_i = ?$

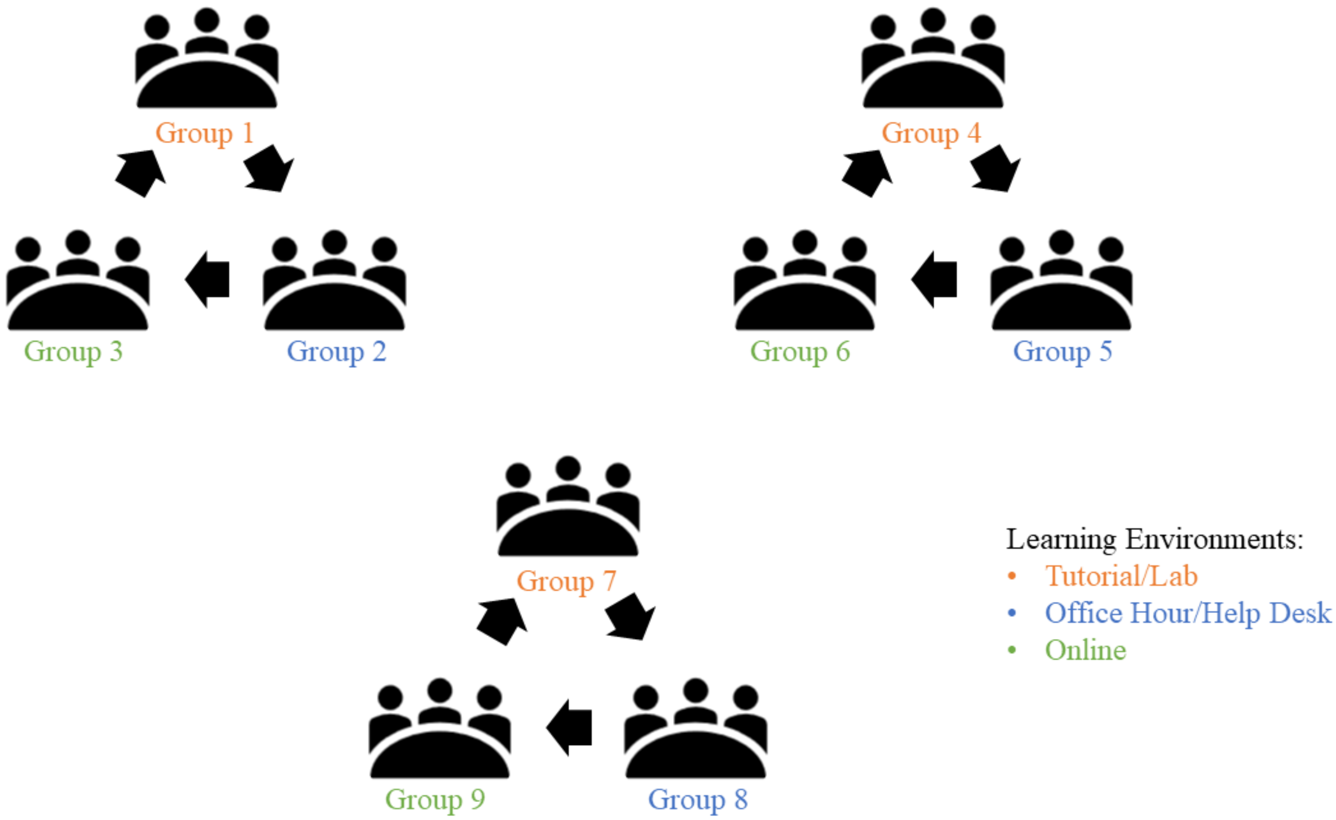
$$
\begin{aligned}
Var(\hat{\beta}_1) &= Var\left[\sum_{i=1}^{n} (y_i - \beta_0) c_i\right] \\
&= \Sigma c_i^2 \, Var(y_i - \beta_0) + 2 \sum\sum_{i \neq j} Cov[(y_i - \beta_0), (y_j - \beta_0)] \\
&= \Sigma c_i^2 \, Var(y_i - \beta_0), \text{ since } Cov[(y_i - \beta_0), (y_j - \beta_0)] = 0 \\
&\qquad \because y_i, y_j \text{ are independent} \\
&= \Sigma c_i^2 \, Var(y_i) \\
&= \sigma^2 \, \Sigma \left(\frac{x_i}{\Sigma x_i^2} \right)^2 \\
&= \sigma^2 \, \frac{\Sigma x_i^2}{(\Sigma x_i^{\cancel{4}2})^2} \\
&= \frac{\sigma^2}{\sum_{i=1}^{n} x_i^2}
\end{aligned}
$$

$\therefore$ statement?

Appendix C

Café Conversation Activity - Instructions

Participants will move around the room in a clockwise manner (shown below). Each table will represent a different learning environment. Therefore, as the activity progresses, each group will have the opportunity to interact with all three environments. The number of groups required for the activity (i.e., three, six, or nine) and the number of participants within each group will depend on the number of participants in attendance.



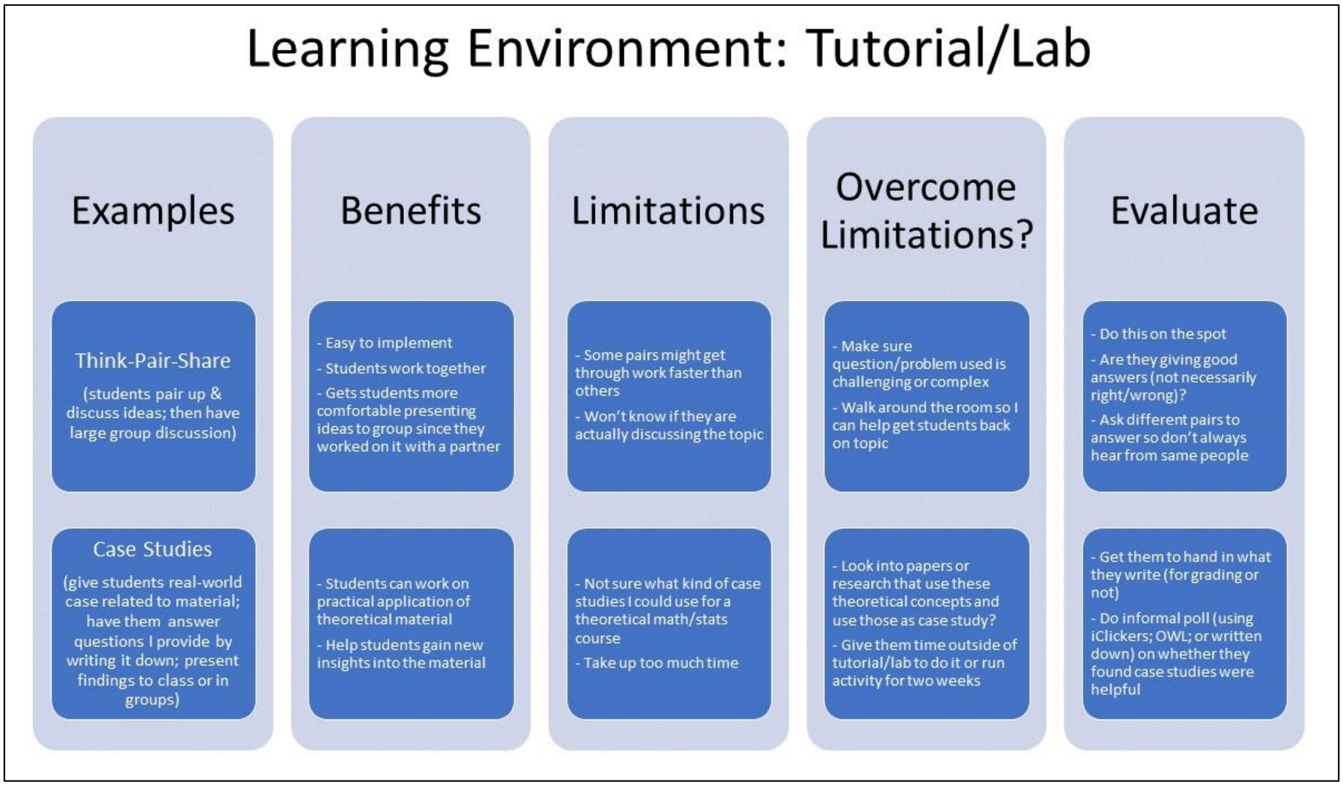
The schedule for the activity is:

☐ Reflect: Active learning activities that participants have seen/experienced (1-2 mins)
☐ Write: Examples of active learning activities with brief explanations (10 mins)
☐ Move tables
☐ Write: Benefits and limitations of active learning activities (10 mins)
☐ Move tables
☐ Write: How you will overcome any limitations and evaluate active learning activities (10 mins)
☐ Move tables
☐ Read and discuss all comments at your table (5 mins)
☐ Debrief activity (15 mins)

Appendix D

Café Conversation Activity - Example

The following image shows an example of a completed chart for the Café Conversation activity. Two active learning activities are provided along with an analysis of how they could be used in a tutorial or lab for math/stats courses.



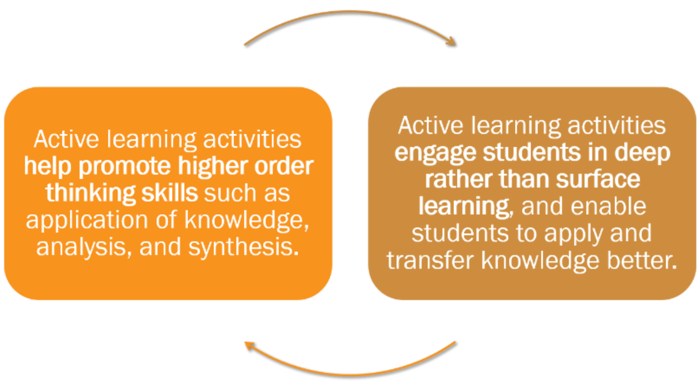
Appendix E

Workshop Takeaway Handout

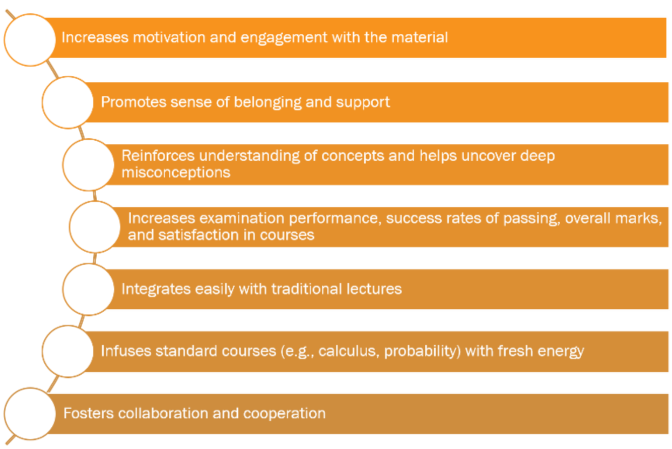
Active Learning in Math & Stats

C. Uggenti, Department of Statistical and Actuarial Sciences, Western University, 2021

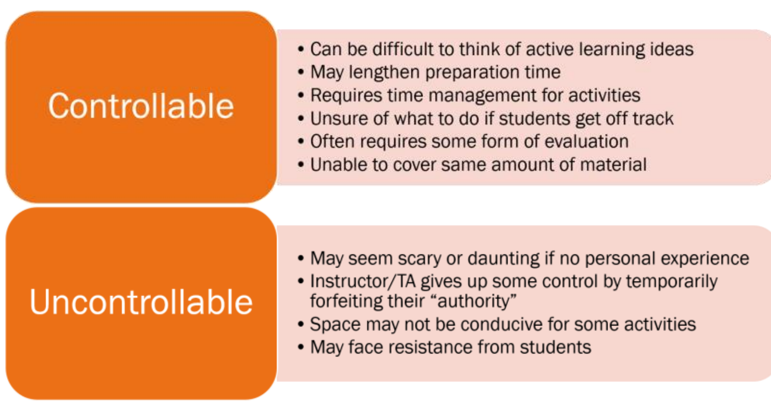
What is Active Learning?



Benefits:



Limitations:



Helpful Websites:

- https://www.queensu.ca/teachingandlearning/modules/active/index.html
- https://teaching.uwo.ca/teaching/learning/active-learning.html
- https://cft.vanderbilt.edu/guides-sub-pages/cats/
- https://www.mghihp.edu/faculty-staff-faculty-compass-teaching-teaching-strategies/examples-classroom-assessment-techniques

Workshop References:

Bonwell, C. C., & Eison, J. A. (1991). *Active learning: creating excitement in the classroom.* ASHE-ERIC Higher Education Reports. Washington, USA.

Braun, B., Bremser, P., Duval, A. M., Lockwood, E., & White, D. (2017). What does active learning mean for mathematicians? *Notices of the American Mathematical Society*, *62*(2), 124-129.

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance is science, engineering, and mathematics. *Proceedings of the National Academy of Sciences*, *111*(23), 8410-8415.

Garfield, J. (1993). Teaching statistics using small-group cooperative learning. *Journal of Statistics Education*, *1*(1), 1-9.

Keeler, C. M., & Steinhorst, R. K. (1995). Using small groups to promote active learning in the introductory statistics course: A report from the field. *Journal of Statistics Education*, *3*(2), 1-9.

Kerrigan, J. (2018). Active learning strategies for the mathematics classroom. *College Teaching*, *66*(1), 35-36.

Prins, S. C. B. (2009). Student-centred instruction in a theoretical statistics course. *Journal of Statistics Education*, *17*(3), 1-12.

Rosenthal, J. S. (1995). Active learning strategies in advanced mathematics classes. *Studies in Higher Education*, *20*(2), 223-228.

References

Angelo, T. A. (1995). Classroom assessment for critical thinking. *Teaching of psychology, 22*(1), 6-7.

Braun, B., Bremser, P., Duval, A. M., Lockwood, E., & White, D. (2017). What does active learning mean for mathematicians? *Notices of the American Mathematical Society, 62*(2), 124-129.

Crowe, E. (2019). *Understanding factors that influence STEM graduate student teaching assistant buy-in to pedagogical training* [Unpublished master's thesis]. The Pennsylvania State University.

Dimitrov, N., Meadows, K., Kustra, E., Ackerson, T., Prada, L., Baker, N., Boulos, P., McIntyre, G., & Potter, M. K. (2013). Assessing graduate teaching development programs for impact on future faculty. *Toronto: Higher Education Quality Council of Ontario*.

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance is science, engineering, and mathematics. *Proceedings of the National Academy of Sciences*, *111*(23), 8410-8415.

Gardner, G. E., & Jones, M. G. (2011). Pedagogical preparation of the science graduate teaching assistant: Challenges and implications. *Science Educator*, *20*(2), 31-41.

Garfield, J. (1993). Teaching statistics using small-group cooperative learning. *Journal of Statistics Education, 1*(1), 1-9.

Gelman, A. (2005). A course on teaching statistics at the university level. *The American Statistician, 59*(1), 4-7.

Justice, N., Zieffler, A., & Garfield, J. (2017). Statistics graduate teaching assistants' beliefs, practices and preparation for teaching introductory statistics. *Statistics Education Research Journal, 16*(1), 294-319.

Keeler, C. M., & Steinhorst, R. K. (1995). Using small groups to promote active learning in the introductory statistics course: A report from the field. *Journal of Statistics Education*, *3*(2), 1-9.

Kerrigan, J. (2018). Active learning strategies for the mathematics classroom. *College Teaching*, *66*(1), 35-36.

Lang, J. M. (2016). *Small teaching: Everyday lessons from the science of learning*. John Wiley & Sons. San Francisco, USA.

Prins, S. C. B. (2009). Student-centred instruction in a theoretical statistics course. *Journal of Statistics Education*, *17*(3), 1-12.

Rosenthal, J. S. (1995). Active learning strategies in advanced mathematics classes. *Studies in Higher Education*, *20*(2), 223-228.

D.2 Pre-Workshop Survey Questionnaire

The following pages contain the questionnaire that study participants completed before the workshop. It consists of demographic questions and workshop-related questions in mixed format (i.e., multiple-choice questions and short answer questions).

GTA Study in SMSS – Pre-Workshop Survey Questionnaire

Section 1: Demographic Questions

Please answer the following questions about yourself.

1. What is your degree program?

   a. Actuarial Science
   b. Applied Mathematics
   c. Financial Modelling
   d. Mathematics
   e. Statistics

2. What is your program?

   a. Master's
   b. Ph.D.

3. Based on the previous question, what year of study are you in?

   a. Year 1
   b. Year 2
   c. Year 3
   d. Year 4
   e. Year 5+

4. What is your current residency status?

   a. Domestic/Permanent Resident (i.e., Canadian or a permanent resident of Canada)
   b. International

5. To which gender identity do you most identify with?

   a. Male
   b. Female
   c. I self-identify as: (with text box)
   d. Prefer not to say

6. What is your age?

   a. 19 years old or younger
   b. 20-24 years old
   c. 25-29 years old
   d. 30-34 years old
   e. 35 years or older
   f. Prefer not to say

7. Is English your first language?

a. Yes
b. No

8. How much experience do you have working as a Graduate Teaching Assistant (GTA)? Here, an academic term is considered four (4) months.

a. 0 academic terms (no experience)
b. 1-2 academic terms
c. 3-4 academic terms
d. 5+ academic terms

9. Please indicate whether you have had any of the following experiences with different types of professional development activities related to teaching.

	Yes	No
Conference		
Short workshops (1-5 hours) e.g., Future Prof Series		
Medium workshops (1-2 days) e.g., TA Day		
Long workshops (3-10 days) e.g., TA Training Program, Advanced Teaching Program		
Summer or semester-long course e.g., SGPS 9500 course		
Other (please specify below)		

10. Other (from Question 9 above, if applicable)

Section 2: Workshop-Related Questions

11. Fill in the blank: As a GTA, I have ________ considered using active learning strategies in labs, tutorials, and/or office hours.

a. Always
b. Often
c. Sometimes
d. Rarely
e. Never
f. Not Applicable

12. Fill in the blank: I am ________ with active learning strategies used in my discipline.

   a. Extremely familiar
   b. Very familiar
   c. Moderately familiar
   d. Slightly familiar
   e. Not at all familiar

13. If you did not answer "Not familiar at all" for the previous question, please answer the following:

   i. What do you know about active learning strategies?

   ii. What would limit you from using active learning strategies in your role as a GTA?

14. Fill in the blank: I feel ________ about my GTA roles and duties.

   a. Very excited
   b. Moderately excited
   c. Indifferent
   d. Moderately nervous
   e. Very nervous

15. As a GTA, I feel comfortable incorporating active learning strategies in labs, tutorials, and/or office hours.

   a. Yes
   b. No
   c. Not Applicable

16. If you responded "Yes" to the previous question, what is an example of an active learning strategy you would use? If you responded "No", please explain why you are not comfortable incorporating these strategies.

17. Fill in the blank: I am ________ in learning more about active learning strategies that I can use as a GTA in my discipline.

   a. Very interested
   b. Somewhat interested
   c. Neither interested nor uninterested
   d. Somewhat uninterested
   e. Very uninterested

18. Fill in the blank: In my discipline, it is ________ to use active learning strategies in undergraduate courses.

   a. Very important
   b. Somewhat important
   c. Neither important nor unimportant
   d. Of low importance
   e. Not important at all

19. I know where to find resources on active learning strategies that can be used in my discipline.

   a. Strongly agree
   b. Agree
   c. Neither agree nor disagree
   d. Disagree
   e. Strongly disagree

Section 3: Final Comments

20. Please add any other comments you have below.

D.3 Post-Workshop Survey Questionnaire

The following pages contain the questionnaire that study participants completed after the workshop. It consists of workshop-related questions in mixed format (i.e., multiple-choice questions and short answer questions).

GTA Study in SMSS – Post-Workshop Survey Questionnaire

Section 1: Workshop-Related Questions

1. Fill in the blank: As a Graduate Teaching Assistant (GTA), I have ________ considered using active learning strategies in labs, tutorials, and/or office hours.

a. Always
b. Often
c. Sometimes
d. Rarely
e. Never
f. Not Applicable

2. Fill in the blank: I am ________ with active learning strategies used in my discipline.

a. Extremely familiar
b. Very familiar
c. Moderately familiar
d. Slightly familiar
e. Not at all familiar

3. If you <u>did not</u> answer "Not familiar at all" for the previous question, please answer the following:

i. What do you know about active learning strategies?

ii. What would limit you from using active learning strategies in your role as a GTA?

4. Fill in the blank: I feel ________ about my GTA roles and duties.

a. Very excited
b. Moderately excited
c. Indifferent
d. Moderately nervous
e. Very nervous

5. As a GTA, I feel comfortable incorporating active learning strategies in labs, tutorials, and/or office hours.

a. Yes
b. No
c. Not Applicable

6. If you responded "Yes" to the previous question, what is an example of an active learning strategy you would use? If you responded "No", please explain why you are not comfortable incorporating these strategies.

7. Fill in the blank: I am ________ in learning more about active learning strategies that I can use as a GTA in my discipline.

   a. Very interested
   b. Somewhat interested
   c. Neither interested nor uninterested
   d. Somewhat uninterested
   e. Very uninterested

8. Fill in the blank: In my discipline, it is ________ to use active learning strategies in undergraduate courses.

   a. Very important
   b. Somewhat important
   c. Neither important nor unimportant
   d. Of low importance
   e. Not important at all

9. I know where to find resources on active learning strategies that can be used in my discipline.

   a. Strongly agree
   b. Agree
   c. Neither agree nor disagree
   d. Disagree
   e. Strongly disagree

10. Fill in the blank: After completing this workshop, I feel __________ about completing my GTA assignment.

   a. More excited
   b. About the same
   c. Less excited

11. Fill in the blank: This workshop has had a __________ impact on my perception of teaching as a GTA in my discipline.

   a. Very positive
   b. Positive
   c. Neither positive nor negative
   d. Negative
   e. Very negative

Section 2: Final Comments

12. Please add any other comments you have about the workshop, active learning strategies, or your experience as a GTA in the space below.

Bibliography

Allignol, A., Schumacher, M., and Beyersmann, J. (2011). Empirical transition matrix of multi-state models: The **etm** package. *Journal of Statistical Software*, 38(4):1–15.

Alvares, D., Lázaro, E., Gómez-Rubio, V., and Armero, C. (2021). Bayesian survival analysis with BUGS. *Statistics in Medicine*, 40(12):2975–3020.

ASA. American Statistical Association, *Education*. Retrieved December 5, 2020, from https://www.amstat.org/asa/education/home.aspx?hkey=81d9c142-50f7-4706-b8ce-7a2edcef2da4.

Austin, P. C. (2017). A tutorial on multilevel survival analysis: Methods, models and applications. *International Statistical Review*, 85(2):185–203.

Birch, J. B. and Morgan, J. (2005). TA Training at Virginia Tech: A Stepwise Progression. *The American Statistician*, 59(1):14–18.

Boman, J. S. (2008). Outcomes of a graduate teaching assistant training program. *ProQuest Dissertations & Theses*.

Bonwell, C. C. and Eison, J. A. (1991). *Active Learning: Creating Excitement in the Classroom*. ASHE-ERIC Higher Education Reports. Washington, USA.

Braun, B., Bremser, P., Duval, A. M., Lockwood, E., and White, D. (2017). What does active learning mean for mathematicians? *Notices of the American Mathematical Society*, 62(2):127–129.

Braxton, J. M., Jones, W. A., Hirschy, A. S., and Hartley III, H. V. (2008). The role of active learning in college student persistence. *New Directions for Teaching and Learning*, 2008(115):71–83.

Carlson, K. A. and Winquist, J. R. (2011). Evaluating an Active Learning Approach to Teaching Introductory Statistics: A Classroom Workbook Approach. *Journal of Statistics Education*, 19(1):1–23.

Carver, R., Everson, M., Gabrosek, J., Horton, N., Lock, R., Mocko, M., Rossman, A., Rowell, G. H., Velleman, P., Witmer, J., and Wood, B. (2016). Guidelines for Assessment and Instruction in Statistics Education (GAISE) College Report 2016. *ASA*. Retrieved February 12, 2022, from https://www.amstat.org/education/guidelines-for-assessment-and-instruction-in-statistics-education-(gaise)-reports.

Colosi, L. and Dunifon, R. (2006). *What's the Difference? "Post then Pre" & "Pre then Post"*. Retrieved December 5, 2020, from http://www.healthymarriageinfo.org/wp-content/uploads/2018/05/What-s-20the-20Difference-20Post-20then-20Pre-20and-20Pre-20then-20Post.pdf.

Cook, R. J. and Lawless, J. F. (2018). *Multistate Models for the Analysis of Life History Data*. Boca Raton, Florida: CRC Press.

Cox, D. R. (1972). Regression Models and Life-Tables. *Journal of the Royal Statistical Society: Series B (Methodological)*, 34(2):187–202.

Crowe, E. (2019). *Understanding Factors That Influence STEM graduate student teaching assistant buy-in to pedagogical training* [Unpublished master's thesis]. The Pennsylvania State University.

Dale, E. (1969). *Audio-Visual Methods in Teaching*. Holt, New York.

de Wreede, L. C., Fiocco, M., and Putter, H. (2011). *Mstate*: An *R* Package for the Analysis of Competing Risks and Multi-State Models. *Journal of Statistical Software*, 38(7):1-30. https://www.jstatsoft.org/v38/i07/.

Denwood, M. J. (2016). Runjags: An R package providing interface utilities, model templates, parallel computing methods and additional distributions for MCMC models in JAGS. *Journal of statistical software*, 71(1):1–25.

Dimitrov, N., Meadows, K., Kustra, E., Ackerson, T., Prada, L., Baker, N., Boulos, P., McIntyre, G., and Potter, M. K. (2013). Assessing Graduate Teaching Development Programs for Impact on Future Faculty. *Toronto: Higher Education Quality Council of Ontario*.

Duchateau, L. and Janssen, P. (2007). *The Frailty Model*. Springer Science & Business Media.

Duchateau, L., Janssen, P., Lindsey, P., Legrand, C., Nguti, R., and Sylvester, R. (2002). The shared frailty model and the power for heterogeneity tests in multicenter trials. *Computational Statistics & Data Analysis*, 40(3):603–620.

Ebert-May, D., Brewer, C., and Allred, S. (1997). Innovation in Large Lectures: Teaching for Active Learning. *BioScience*, 47(9):601–607.

Faust, J. L. and Paulson, D. R. (1998). Active Learning in the College Classroom. *Journal on Excellence in College Teaching*, 9(2):3–24.

Felder, R. M. and Brent, R. (1996). Navigating the Bumpy Road to Student-Centered Instruction. *College Teaching*, 44(2):43–47.

Fox, J. and Weisberg, S. (2002). Cox Proportional-Hazards Regression for Survival Data in R. In *An R and S-PLUS Companion to Applied Regression*, pages 1–19. Sage, Thousand Oaks, CA.

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., and Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences*, 111(23):8410–8415.

French, S. and Kennedy, G. (2017). Reassessing the value of university lectures. *Teaching in Higher Education*, 22(6):639–654.

Froelich, A. G., Duckworth, W. M., and Stephenson, W. R. (2005). Training Statistics Teachers at Iowa State University. *The American Statistician*, 59(1):8–10.

Gabry, J., Simpson, D., Vehtari, A., Betancourt, M., and Gelman, A. (2019). Visualization in Bayesian workflow. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 182(2):389–402.

Gardner, G. E. and Jones, M. G. (2011). Pedagogical preparation of the science graduate teaching assistant: Challenges and implications. *Science Educator*, 20(2):31–41.

Garfield, J. (1993). Teaching Statistics Using Small-Group Cooperative Learning. *Journal of Statistics Education*, 1(1):1–9.

Garfield, J. (1995). How Students Learn Statistics. *International Statistical Review / Revue Internationale de Statistique*, 63(1):25–34.

Garfield, J. and Ben-Zvi, D. (2007). How Students Learn Statistics Revisited: A Current Review of Research on Teaching and Learning Statistics. *International Statistical Review*, 75(3):372–396.

Garfield, J. and Everson, M. (2009). Preparing Teachers of Statistics: A Graduate Course for Future Teachers. *Journal of Statistics Education*, 17(2).

Gelman, A. (2005). A Course on Teaching Statistics at the University Level. *The American Statistician*, 59(1):4–7.

Gelman, A. (2006). Prior distributions for variance parameters in hierarchical models. *Bayesian analysis*, 1(3):515–533.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., and Rubin, D. B. (2014a). *Bayesian Data Analysis*. Chapman and Hall/CRC, third edition.

Gelman, A., Hwang, J., and Vehtari, A. (2014b). Understanding predictive information criteria for Bayesian models. *Statistics and computing*, 24(6):997–1016.

Gelman, A. and Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, 7(4):457–472.

Gilmore, J., Maher, M. A., Feldon, D. F., and Timmerman, B. (2014). Exploration of factors related to the development of science, technology, engineering, and mathematics graduate teaching assistants' teaching orientations. *Studies in Higher Education*, 39(10):1910–1928.

Graham, I. D., Logan, J., Harrison, M. B., Straus, S. E., Tetroe, J., Caswell, W., and Robinson, N. (2006). Lost in knowledge translation: Time for a map? *Journal of Continuing Education in the Health Professions*, 26(1):13–24.

Grambsch, P. M. and Therneau, T. M. (1994). Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika*, 81(3):515–526.

Green, J. L. (2010). Teaching Highs and Lows: Exploring University Teaching Assistants' Experiences. *Statistics Education Research Journal*, 9(2):108–122.

Grimshaw, J. M., Eccles, M. P., Lavis, J. N., Hill, S. J., and Squires, J. E. (2012). Knowledge translation of research findings. *Implementation Science*, 7(1):1–17.

Harkness, W. and Rosenberger, J. (2005). Training Graduate Students at Penn State University in Teaching Statistics. *The American Statistician*, 59(1):11–13.

Harrell, F. E. (2015). *Regression Modeling Strategies: With Applications to Linear Models, Logistic and Ordinal Regression, and Survival Analysis*. Springer Series in Statistics. Springer International Publishing, Cham.

Hertz-Picciotto, I. and Rockhill, B. (1997). Validity and Efficiency of Approximation Methods for Tied Survival Times in Cox Regression. *Biometrics*, 53(3):1151–1156.

Howard, G. S. (1980). Response-Shift Bias: A Problem in Evaluating Interventions with Pre/Post Self-Reports. *Evaluation Review*, 4(1):93–106.

Ibrahim, J. G., Chen, M.-H., and Sinha, D. (2001). *Bayesian Survival Analysis*. Springer Science & Business Media.

Jackson, C. H. (2016). Flexsurv: A platform for parametric survival modeling in R. *Journal of Statistical Software*, 70.

Johns, R. (2010). Likert items and scales. *Survey Question Bank: Methods Fact Sheet 1*.

Justice, N. (2020). Preparing Graduate Students to Teach Statistics: A Review of Research and Ten Practical Recommendations. *Journal of Statistics Education*, 28(3):334–343.

Justice, N., Zieffler, A., and Garfield, J. (2017). Statistics Graduate Teaching Assistants' Beliefs, Practices and Preparation for Teaching Introductory Statistics. *Statistics Education Research Journal*, 16(1).

Kalbfleisch, J. D. and Prentice, R. L. (1973). Marginal Likelihoods Based on Cox's Regression and Life Model. *Biometrika*, 60(2):267–278.

Kanevsky, L. (2016). Assessing Students' Perceptions of the Effectiveness of Instructional Interventions with Post-Pre Surveys. Retrieved February 5, 2022, from https://www.sfu.ca/content/dam/sfu/istld/documents/Post-Pre_Kanevsky_handout.pdf.

Keeler, C. M. and Steinhorst, R. K. (1995). Using Small Groups to Promote Active Learning in the Introductory Statistics Course: A Report from the Field. *Journal of Statistics Education*, 3(2):1–9.

Kerrigan, J. (2018). Active Learning Strategies for the Mathematics Classroom. *College Teaching*, 66(1):35–36.

Kleinbaum, D. G. and Klein, M. (2010). *Survival Analysis*, volume 3. New York: Springer.

Korsgaard, R., I., Madsen, P., and Jensen, J. (1998). Bayesian inference in the semiparametric log normal frailty model using Gibbs sampling. *Genetics Selection Publication*, 30(3):241–256.

Kothari, A. and Wathen, C. N. (2017). Integrated knowledge translation: Digging deeper, moving forward. *Journal of Epidemiology and Community Health*, 71(6):619–623.

Kruschke, J. (2015). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan*. Academic Press, second edition.

Lawless, J. F. (2011). *Statistical Models and Methods for Lifetime Data*, volume 362. John Wiley & Sons.

Lindgren, F. and Rue, H. (2015). Bayesian Spatial Modelling with R-INLA. *Journal of Statistical Software*, 63(19):1-25. https://doi.org/10.18637/jss.v063.i19.

Lopatto, D. (2007). Undergraduate Research Experiences Support Science Career Decisions and Active Learning. *CBE—Life Sciences Education*, 6(4):297–306.

McFayden, C. B., Johnston, L. M., Woolford, D. G., George, C., Boychuk, D., Johnston, D., Wotton, B. M., and Johnston, J. M. (2022). A conceptual framework for knowledge exchange in a wildland fire research and practice context [submitted].

Meira-Machado, L., de Uña-Álvarez, J., Cadarso-Suárez, C., and Andersen, P. K. (2009). Multi-state models for the analysis of time-to-event data. *Statistical Methods in Medical Research*, 18(2):195–222.

Mello, D. and Less, C. A. (2013). Effectiveness of active learning in the arts and sciences. *Humanities Department Faculty Publications & Research*.

Michael, J. (2006). Where's the evidence that active learning works? *Advances in Physiology Education*, 30(4):159–167.

Montgomery, D. C., Peck, E. A., and Vining, G. G. (2012). *Introduction to Linear Regression Analysis*. John Wiley & Sons, Hoboken, New Jersey, fifth edition.

Moore, D. S. (2005). Preparing Graduate Students to Teach Statistics: Introduction. *The American Statistician*, 59(1):1–3.

Morin, A. A., Albert-Green, A., Woolford, D. G., and Martell, D. L. (2015). The use of survival analysis methods to model the control time of forest fires in Ontario, Canada. *International Journal of Wildland Fire*, 24(7):964–973.

Morin, A. A., Albert-Green, A., Woolford, D. G., and Martell, D. L. (2019). Frailty models for the control time of wildland fires in the former intensive fire management zone of Ontario, Canada. *Journal of Environmental Statistics*, 9(5):1–16.

Nathoo, F. S. and Dean, C. B. (2008). Spatial Multistate Transitional Models for Longitudinal Event Data. *Biometrics*, 64(1):271–279.

Natural Resources Canada (2021a). *Cost of Wildland Fire Protection*. Retrieved March 30, 2022, from https://www.nrcan.gc.ca/climate-change/impacts-adaptations/climate-change-impacts-forests/forest-change-indicators/cost-fire-protection/17783.

Natural Resources Canada (2021b). *Forest Fire Danger Rating Tool*. Retrieved April 4, 2022, from https://www.nrcan.gc.ca/our-natural-resources/forests-forestry/wildland-fires-insects-disturban/forest-fire-danger-rating-tool/14470.

Ntzoufras, I. (2009). *Bayesian Modeling Using WinBUGS*, volume 698 of *Wiley Series in Computational Statistics*. John Wiley & Sons.

OMNR (2004). *Forest Fire Management Strategy for Ontario*. Queen's Printer for Ontario, Toronto.

OMNRF (2014a). *Forest Fire Management*. Retrieved December 5, 2020, from https://www.ontario.ca/page/forest-fire-management.

OMNRF (2014b). *Wildland Fire Management Strategy*. Queen's Printer for Ontario, Toronto.

O'Neill, G. and McMahon, T. (2005). Student-centred learning: What does it mean for students and lecturers? *Emerging Issues in the Practice of University Learning and Teaching*. Dublin: AISHE.

Pentecost, T. C., Langdon, L. S., Asirvatham, M., Robus, H., and Parson, R. (2012). Graduate teaching assistant training that fosters student-centered instruction and professional development. *Journal of College Science Technology*, pages 68–75.

Plummer, M. (2003). JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. *3rd International Workshop on Distributed Statistical Computing (DSC 2003); Vienna, Austria*, 124(4):1–10.

Potter, M. K. and Kustra, E. (2011). The Relationship between Scholarly Teaching and SoTL: Models, Distinctions, and Clarifications. *International Journal for the Scholarship of Teaching and Learning*, 5(1).

Prince, M. (2004). Does Active Learning Work? A Review of the Research. *Journal of Engineering Education*, 93(3):223–231.

Prins, S. C. B. (2009). Student-Centered Instruction in a Theoretical Statistics Course. *Journal of Statistics Education*, 17(3):1–12.

Putter, H., Fiocco, M., and Geskus, R. B. (2007). Tutorial in Biostatistics: Competing Risks and Multi-State Models. *Statistics in Medicine*, 26(11):2389–2430.

R Core Team (2021). R: A Language and Environment for Statistical Computing. *R Foundation for Statistical Computing*, Vienna, Austria. https://www.R-project.org/.

Reeves, T. D., Marbach-Ad, G., Miller, K. R., Ridgway, J., Gardner, G. E., Schussler, E. E., and Wischusen, E. W. (2016). A Conceptual Framework for Graduate Teaching Assistant Professional Development Evaluation and Research. *CBE—Life Sciences Education*, 15(2):es2.

Rivera, S. (2018). A Summer Institute for STEM Graduate Teaching Assistants: Exploring Teaching Perceptions. *Journal of College Science Teaching*, 48(2):28–32.

Rondeau, V. (2010). Statistical models for recurrent events and death: Application to cancer events. *Mathematical and Computer Modelling*, 52(7-8):949–955.

Rosenthal, J. S. (1995). Active learning strategies in advanced mathematics classes. *Studies in Higher Education*, 20(2):223–228.

Rue, H., Martino, S., and Chopin, N. (2009). Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 71(2):319–392.

Salkind, N. J. (2010). *Encyclopedia of Research Design*. SAGE Publications, Inc., Thousand Oaks, CA.

Skeff, K. M., Stratos, G. A., and Bergen, M. R. (1992). Evaluation of a Medical Faculty Development Program: A Comparison of Traditional Pre/Post and Retrospective Pre/Post Self-Assessment Ratings. *Evaluation & the Health Professions*, 15(3):350–366.

SSC. Statistical Society of Canada, *Education*. Retrieved December 5, 2020, from https://ssc.ca/en/education.

Stocks, B. and Martell, D. L. (2016). Forest fire management expenditures in Canada: 1970–2013. *The Forestry Chronicle*, 92(03):298–306.

Sun, C. (2013). Bivariate extreme value modeling of wildland fire area and duration. *Forest Science*, 59(6):649–660.

Taylor, S. W., Woolford, D. G., Dean, C. B., and Martell, D. L. (2013). Wildfire prediction to inform management: Statistical science challenges. *Statistical Science*, 28(4):586–615.

Therneau, T. M. (2020). A package for survival analysis in R. *R package version 3.2-7*. https://CRAN.R-project.org/package=survival.

Therneau, T. M. and Grambsch, P. M. (2000). *Modeling Survival Data: Extending the Cox Model*. Springer.

Trigwell, K., Prosser, M., and Ginns, P. (2005). Phenomenographic pedagogy and a revised *Approaches to teaching inventory. Higher Education Research & Development*, 24(4):349–360.

Van Wagner, C. E. (1987). Development and Structure of the Canadian Forest Fire Weather Index System. *Canadian Forest Service*.

Vehtari, A., Gabry, J., Magnusson, M., Yao, Y., Bürkner, P.-C., Paananen, T., and

Gelman, A. (2020). Loo: Efficient Leave-One-Out Cross-Validation and WAIC for Bayesian Models. *R package version 2.4.1*. https://mc-stan.org/loo/.

Vehtari, A., Gelman, A., and Gabry, J. (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing*, 27(5):1413–1432.

Vehtari, A., Simpson, D., Gelman, A., Yao, Y., and Gabry, J. (2021). Pareto Smoothed Importance Sampling. *arXiv:1507.02646 [stat.CO]*.

Venables, W. N. and Ripley, B. D. (2002). *Modern Applied Statistics with S*. Springer, New York, fourth edition.

Wachowicz, M. and Chrisman, N. R. (2012). *The Added Value of Scientific Networking: Perspectives from the GEOIDE Network Members 1998-2012*.

Waldrop, M. M. (2015). The science of teaching science. *Nature*, 523(7560):272.

Ward, V., Smith, S., House, A., and Hamer, S. (2012). Exploring knowledge exchange: A useful framework for practice and policy. *Social Science & Medicine*, 74(3):297–304.

Watanabe, S. and Opper, M. (2010). Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory. *Journal of Machine Learning Research*, 11(12).

Williams, C., Lewsey, J. D., Briggs, A. H., and Mackay, D. F. (2017). Cost-effectiveness analysis in R using a multi-state modeling survival analysis framework: A tutorial. *Medical Decision Making*, 37(4):340–352.

Williams, L. S. (1991). The effects of a comprehensive teaching assistant training program on teaching anxiety and effectiveness. *Research in Higher Education*, 32(5):585–598.

Wood, S. N. (2006). *Generalized Additive Models: An Introduction with R*. CRC Press.

Wotton, B. M. (2009). Interpreting and using outputs from the Canadian Forest Fire Danger Rating System in research applications. *Environmental and Ecological Statistics*, 16(2):107–131.

Wotton, B. M. and Martell, D. L. (2005). A lightning fire occurrence model for Ontario. *Canadian Journal of Forest Research*, 35(6):1389–1401.

Wright, G. B. (2011). Student-Centered Learning in Higher Education. *International Journal of Teaching and Learning in Higher Education*, 23(1):92–97.

Xi, D. D., Dean, C. B., and Taylor, S. W. (2020). Modeling the duration and size of extended attack wildfires as dependent outcomes. *Environmetrics*, 31(5):e2619.

Xi, D. D., Taylor, S. W., Woolford, D. G., and Dean, C. (2019). Statistical Models of Key Components of Wildfire Risk. *Annual Review of Statistics and Its Application*, 6(1):197–222.

Xi, D. D. Z., Dean, C. B., and Taylor, S. W. (2021). Modeling the duration and size of wildfires using joint mixture models. *Environmetrics*, 32(6):e2685.

Yazedjian, A. and Kolkhorst, B. B. (2007). Implementing Small-Group Activities in Large Lecture Classes. *College Teaching*, 55(4):164–169.

Zakrajsek, T. (2018). Reframing the lecture versus active learning debate: Suggestions for a new way forward. *Education in the Health Professions*, 1(1):1–3.

Curriculum Vitae

Name:	Chelsea Uggenti
Post-Secondary Education and Degrees:	Ph.D. Statistics, Environment and Sustainability Collaborative Specialization, University of Western Ontario, 2017-*present* M.Sc. Statistics, University of Western Ontario, 2016-2017. B.A. Honours Mathematics, Wilfrid Laurier University, 2012-2016.
Awards:	*A.E.R. Donor Scholarship*, Centre for Environment and Sustainability, University of Western Ontario, 2021. *Postgraduate Scholarship - Doctoral Program*, NSERC, University of Western Ontario, 2017-2020. *Travel Award*, Centre for Environment and Sustainability, University of Western Ontario, 2019. *Science International Engagement Fund*, Science International, University of Western Ontario, 2019. *Graduate Student Award (Merit)*, Centre for Environment and Sustainability, University of Western Ontario, 2019. *Best Poster Recipient (Statistics)*, DSAS M.Sc. Day, University of Western Ontario, 2017. *Canada Graduate Scholarship - Master's Program*, NSERC, University of Western Ontario, 2016-2017. *A1 Research & Travel Grant*, MS2Discovery Institute, Wilfrid Laurier University, 2016. *Travel & Equipment Grant*, Women in Science, Wilfrid Laurier University, 2016. *Undergraduate Student Research Award*, NSERC, Wilfrid Laurier University, 2016. *Ken McDowell Mathematics Scholarship*, Wilfrid Laurier University, 2015.

Awards Continued:	*Best Poster Recipient*, Canadian Mathematical Society - SARGC, Statistical Society of Canada, 2015. *Undergraduate Student Research Award*, NSERC, Wilfrid Laurier University, 2015. *Research Assistant Scholarship*, Faculty of Science Student's Association, Wilfrid Laurier University, 2014-2015. *President's Gold Scholarship*, Wilfrid Laurier University, 2012-2016. *Centennial Scholarship*, Wilfrid Laurier University, 2012-2013.
Related Work Experience:	*Teaching Assistant Training Program Instructor*, Centre for Teaching and Learning, University of Western Ontario, 2020-2022. *Sessional Lecturer* - Statistical Science 3843A, University of Western Ontario, Fall 2018 and Fall 2021. *Teaching Assistant*, Statistical Sciences 1024 and 4844/9544, University of Western Ontario, 2016-2022. *R Workshop Instructor & Statistical Consultant*, Western Data Science Solutions, 2018-2021.
Service:	*Co-Organizer & Chair*, Statistical Education Section sponsored session on "Active learning in statistics: where are we now"? Statistical Society of Canada's Annual Meeting, May - June 2022. *Organizer*, EnviroCon 2019 & 2021 Conferences, University of Western Ontario, 2018-2021. *Visiting Research Scholar*, Dr. Meg Krawchuk's Landscape Fire and Conservation Science Research Group, Oregon State University, 2019. *Moderator*, Panel on Overcoming Barriers to Composting, University of Western Ontario, 2018. *Facilitator*, Shad Canada, University of Western Ontario, 2018. *Undergraduate Student Representative (Elect)*, Mathematics Department, Wilfrid Laurier University, 2015-2016.

Peer-Reviewed Publications:

Uggenti, C., and McCluskey, C.C. (2018). Global stability for infectious disease models that include immigration of infected individuals and delay in the incidence. *Electronic Journal of Differential Equations, 2018*(64), 1-14.

Non-Peer-Reviewed Publications:

Demand, M., Ng, K., and **Uggenti, C.** (2021, March 15). *Fighting Fire with Fire.* Alternatives Journal.
https://www.alternativesjournal.ca/science-research/fighting-fire-with-fire/

Invited Presentations:

Uggenti, C., Woolford, D.G., and McFayden, C. (2015). Exploring the impact of restricted fire zones on the risk of people-caused forest fires in Ontario. *Ontario Ministry of Natural Resources and Forestry Collaborative Research Agreement and Decision Support Systems Meeting.* University of Toronto. Toronto, ON. Aug. 2015.

Contributed Presentations:

Uggenti, C., Woolford, D.G., and Dean, C. B. Characterizing the lifetime phases of wildland fires from the Sioux Lookout District in Ontario by utilizing mixed effects multistate modelling techniques. *Annual Meeting of the Statistical Society of Canada.* Online. June 2021.

Uggenti, C. Whose dictionary is it anyway? *Centre for Teaching and Learning's Own Your Future: May Conference on Teaching.* University of Western Ontario. London, ON. May 2021.

Uggenti, C. (2021). A statistician's experience of getting dirt on her boots. *EnviroCon Conference.* University of Western Ontario. London, ON. March 2021.

Uggenti, C., and McCluskey, C.C. (2016). Investigating the stability of disease models with temporary immunity. *Canadian Undergraduate Mathematics Conference.* University of Victoria. Victoria, BC. July 2016.

Uggenti, C., Woolford, D.G., and McFayden, C. (2015). Exploring the impact of restricted fire zones on the risk of people-caused forest fires in Ontario. *Annual Meeting of the Statistical Society of Canada.* Dalhousie University. Halifax, NS. June 2015.

Poster Presentations:

Uggenti, C. Active learning in Math & Stats: a graduate teaching assistant (GTA) training and development program. *United States Conference on Teaching Statistics.* Online. June-July 2021.

Uggenti, C., Woolford, D.G., and Dean, C.B. (2019). Characterizing the lifetime stages of wildland fires. *Wildland Fire Canada Conference.* Ottawa, ON. Nov. 2019.

Uggenti, C., Woolford, D.G., and Dean, C.B. (2017). Investigating fire season lengthening in Alberta, Canada. *Master's of Science Day.* Department of Statistical and Actuarial Sciences, University of Western Ontario. London, ON. July 2017.

Uggenti, C., Stacey, A., Woolford, D.G., and McFayden, C. (2017). The impact of fire bans on recreation and resident-caused wildland fires in Ontario's Restricted Fire Zones. *Fallona Interdisciplinary Showcase.* University of Western Ontario. London, ON. Jan. 2017.

Uggenti, C., and Woolford, D.G. (2015). Exploratory data analysis on fire weather variables to observe bias and variability. *Statistical Society of Canada's Student Conference.* Dalhousie University. Halifax, NS. June 2015.