BAYESIAN HIERARCHICAL MODELLING OF SPRUCE BUDWORM DEVELOPMENT

BAYESIAN HIERARCHICAL MODELLING OF SPRUCE BUDWORM DEVELOPMENT

by KALA STUDENS, B.Math

A Thesis Submitted to the School of Graduate Studies in Partial Fulfilment of the Requirements for the Degree Master of Science

McMaster University MASTER OF SCIENCE (2021) Hamilton, Ontario (Statistics)

TITLE: Bayesian Hierarchical Modelling of Spruce Budworm Development

AUTHOR: Kala Studens, B.Math (University of Waterloo)

SUPERVISOR: Professor B. Bolker

NUMBER OF PAGES x, 51

Lay Abstract

Managing forest pests requires an understanding of their life cycle and the ability to predict when life events will occur. It is also important to know whether insects from different parts of the country will develop in the same way. Many models have been created to capture the relationship between temperature and development; this study aims to implement one of these models in a Bayesian framework to get a clearer picture of uncertainty around parameter estimates and predictions. The model was fitted to data from the larval stages of several spruce budworm colonies, each collected from a different location across Canada. These colonies were reared in laboratory conditions, under constant temperature regimes. Estimates of insect development were compared across colonies and across developmental stages.

Abstract

The management of destructive forest pests such as the spruce budworm relies on accurate modelling of their development. Predicting the timing of specific events in the life cycle is crucial for pest control tactics and for modelling the landscape-scale dispersal of the insect. This thesis implements a Bayesian hierarchical thermal response model for the larval stages of the spruce budworm. The model was fitted to data collected from a laboratory rearing experiment on wild spruce budworm colonies collected from locations across Canada and on a fully lab-reared colony. The results were compared across developmental stages and geographic origins. The Bayesian model was implemented with the non-linear, temperature-dependent development rate curve outlined in Schoolfield et al. (1981) and the framework in Régnière et al. (2012) for individual variation and interval censored data. Posterior samples were obtained and a quadratic relationship was observed between developmental stage and an intercept parameter of the development curve. A second model was fitted to the data incorporating this structure. Distributions of development rate estimates at each rearing temperature were obtained from each posterior sample and it was observed that the lab-reared colony developed more quickly than the wild colonies. In future work, the posterior samples can be used to generate simulated populations for prediction, with uncertainty fully propagated throughout.

Acknowledgements

I would first like to thank my supervisor, Ben Bolker, for his guidance and expertise in helping me to complete this thesis. This process was an invaluable learning experience and I thank you for all that you taught me along the way.

I would also like to thank my colleagues at the Great Lakes Forestry Centre. Particularly, my supervisor Jean-Noel Candau, for his mentorship and his advocacy for the furthering of my education. Thank you for giving me this opportunity and for your support throughout. Also to Ashlyn Wardlaw and Kerry Perrault, whose countless hours and diligence in the lab resulted in this extensive dataset. Thank you both for your friendship and encouragement.

I would like to thank my parents for their love and support, and for always being there to listen and to offer sound advice. Thank you for all of the meals, conversations, saunas and skiing adventures that make life so enjoyable.

And to Peter, thank you for helping me talk through my ideas and giving excellent advice, and for always being there for me.

Contents

List of Figures

List of Abbreviations and Symbols

CDF	Cumulative Distribution Function
ESS	Effective Sample Size
HMC	Hamiltonian Monte Carlo
MCMC	Markov Chain Monte Carlo
NUTS	No-U-Turn Sampler

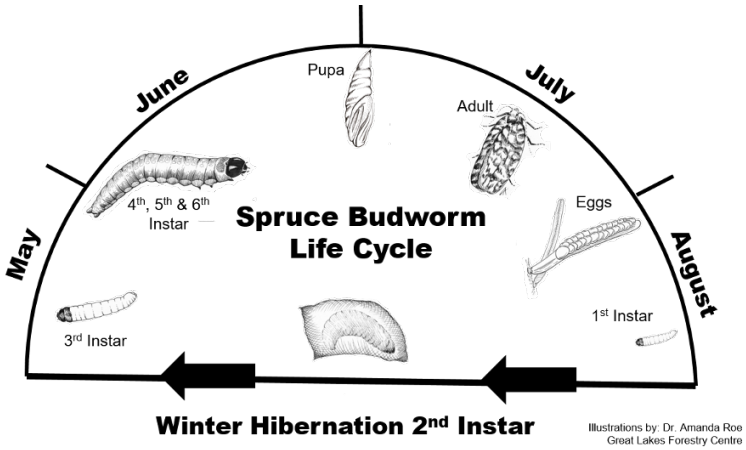
Chapter 1

Introduction

1.1 Background

The eastern spruce budworm, *Choristoneura fumiferana*, is an insect species native to North America. During periodic outbreaks, its populations have caused extensive damage to Canadian forests by defoliation, especially of white spruce and balsam fir trees. The management of such forest pests relies on an understanding of the pest's life cycle and the factors that affect it (Candau et al., 2019). For some pest management tactics, such as the spraying of insecticide, it is crucial to time application accurately in order to target the proper development stage. An illustration of the spruce budworm's life cycle is shown in Figure 1.1. The spruce budworm is univoltine, producing offspring once per year. The insect's larval stages are referred to as instars. Upon emergence from eggs, the first instar larvae disperse and form hibernacula in which they spend the duration of the winter and moult into second instar larvae. Second instar larvae emerge from these hibernacula in the early spring, at which point they begin to feed on old needles and other nutrition that is available prior to new bud growth. In the third to sixth larval instars, larvae feed on new bud growth; these larval stages are delineated by the shedding of head capsules. The synchrony of these larval stages with new bud growth is crucial for survival; understanding the timing of each phenomenon is necessary in order to predict the impact of climate change on the species' behaviour. The subsequent pupal stage is typically reached mid-summer, and moths emerge several days later. Spruce budworm moths have been shown to disperse across great distances, both by flying and being carried by prevailing winds. Understanding the timing of adult emergence is important for modelling the species' landscape-scale

Figure 1.1: Life cycle of the spruce budworm



dispersal. Development rate models form the basis for an accurate understanding of the phenology of the spruce budworm and other insects. While very accurate for certain regions, current predictive tools such as BioSIM (Régnière et al., 2014), do not take the geographic variation of the species into account.

1.2 Review of Literature

1.2.1 Modelling Insect Development

Temperature is the most important factor in the growth and development of insects and other poikilotherms (Uvarov et al., 1931; Régnière and Powell, 2013). Other factors such as photoperiod, humidity and diet play a role in development, but temperature has the strongest impact and its physiological effect is well understood. Most models of insect development are based on temperature dependence (Damos and Savopoulou-Soultani, 2012). The idea of development summation is common to most temperature-dependent development models. That is, the developmental age a of the subject can be determined by the equation

$$a = \int_0^t r(T_{t'})\, dt' \tag{1.1}$$

where $r(T_t)$ represents development rate at temperature T occurring at time t. This model structure presumes that r is an "instantaneous" development rate; that is, an abrupt shift in temperature immediately changes an organism's development rate. This assumption, while extremely common in the literature, has faced criticism, as insects may develop at different rates in fluctuating conditions

than at constant temperatures (Worner, 1992). In nature, insects experience a wide range of fluctuating temperatures. However, due to the constraints involved in observing insect development in a laboratory setting, much of the data used to fit development rate models comes from insects reared under constant temperature regimes. This likely causes bias in the results of fitted development rate models. Using laboratory data to fit these models also implicitly assumes that organisms exhibit the same behaviour under experimental conditions as they do in the wild (Tanigoshi and Logan, 1979).

Degree-Day Models

According to de Candolle's law of total effective temperature, poikilotherms require an accumulation of thermal units up to a species-specific threshold to complete a developmental event; the threshold can vary genetically or geographically (Damos and Savopoulou-Soultani, 2012). This is expressed by the equation

$$k = D(T - T_b)$$

where k is the thermal threshold, D is the duration of growth, T is the temperature experienced, and T_b is a threshold temperature below which development stops. The units of the parameter k are degree-days: accumulated temperature (degrees) above a temperature threshold. A simple transformation of this equation gives what is referred to as the "degree-day" model

$$r(T) = \frac{1}{D} = \begin{cases} \frac{1}{k}(T - T_b) & \text{if } T > T_b \\ 0 & \text{otherwise} \end{cases}$$

in which development rate $1/D$ is represented as a linear function of temperature T. Estimates of the slope and intercept can be obtained using linear regression, and the underlying parameters T_b and k can be inferred.

Degree-day models rely on the assumption of a linear relationship between development and temperature. Davidson (1944) describes the development rate curve for the eggs of *Drosophila melanogaster* as "S-shaped", indicating departures from linearity at both low and high temperatures. Ikemoto and Takai (2000) acknowledge that degree-days are best used on data in an optimal temperature range, and propose a modified degree-day model wherein part of the model fitting involves finding upper and lower bounds for this optimal range. The typical optimal temperature range excludes many of the temperatures that are often experienced in nature. Thus, in order to

accurately represent development outside an optimal range, nonlinear models should be considered.

Nonlinear Models

The development summation principle described above is the same for nonlinear models, but the development rate function r is no longer simply a linear function of temperature; it can be made to have a more complex or flexible functional form. Many different nonlinear development rate functions have been proposed. These models typically fall into one of two classes: statistical or biophysical. Statistical models are created to mimic the desired behaviour of the development curve by engineering the development rate formula to have certain geometric features, while biophysical models are mechanistic representations of a physiological process (Rebaudo and Rabhi, 2018). While one of the main benefits of statistical models is graphical convenience, their parameters may also have biological interpretations.

An early non-linear statistical phenology model was the sigmoidal response proposed by Davidson (1944), based on the same response function Van't Hoff 1884 and Arrhenius 1915, developed for modelling chemical reactions. In fitting this curve Davidson (1944) omitted observations above a temperature threshold where the development began to slow, since the sigmoidal curve is not suited to capture such behaviour. Stinner et al. (1974) proposed a sigmoid curve reflected at the optimal temperature to capture the decreasing development rate at high temperatures, imposing symmetry in the response curve. Other development rate curves with the same feature are the catenary model (Janisch, 1932) and the truncated Gaussian model (Taylor, 1981). While providing an improvement in goodness of fit to the original sigmoidal model, symmetrical models still fail to fit thermal response data well. This issue led to the development of asymmetric, unimodal response curves (Régnière and Powell, 2013).

One such model, developed by Logan et al. (1976), combined the exponentially increasing shape of the observations below the optimal temperature with a quickly decreasing function representing a sharp decline in development until a lethal temperature is reached where zero development occurs. These two phases were combined by matching asymptotes. Unlike many that came before it, this model contains parameters with some biological interpretation. This is a positive feature for understanding the model parameters themselves, making it more intuitive to compare fits across stages or populations. Interpretable parameters are also useful for model fitting; optimization is helped by good starting parameter values and scientists with prior knowledge about insect behaviour are likely

able to provide reasonable estimates. Hilbert and Logan (1983) used the same matched asymptote method as Logan et al. (1976), replacing the original exponential curve of Phase I with Holling's Type III sigmoid equation (Holling, 1965). The alternative sigmoid curve was proposed as an improvement over the original Phase I shape, which only approaches zero asymptotically. Hilbert and Logan (1983) argued that the development curve should feature a lower temperature threshold, in addition to the upper temperature threshold that exists in the model. Lactin et al. (1995) modified the Logan et al. (1976) model, by adding an intercept term and removing an amplitude parameter they claimed was redundant. Brière et al. (1999) proposed a new model with fewer parameters, all with biological interpretability. However, they showed that the original Logan et al. (1976) model provides a better fit to the data than the model with fewer parameters.

A biophysical model featuring both a linear relationship between temperature and development in an optimal intermediate temperature range and a non-linear dampening of development at both high and low temperatures was proposed by Sharpe and DeMichele (1977). This model was based on work by Johnson and Lewin (1946) and Hultin (1955), who introduced models incorporating high and low temperature inhibition of development, respectively. Sharpe and DeMichele (1977) is based upon the assumption that an organism's development relies on a single control enzyme which can exist in one of three reversible states: an active state in which development occurs, and low and high temperature states in which no development occurs. It is also assumed that development is proportional to the concentration of enzymes in each state. Each parameter in this model represents a feature of the theorized chemical mechanism of development, giving them each interpretability.

Schoolfield et al. (1981) claim that Sharpe and DeMichele (1977) is unsuitable for nonlinear regression, due to high correlation between the parameter estimators. Based on Hultin (1955), the new formulation improves the model's suitability for regression by re-defining three of the parameters in terms of more biologically and geometrically interpretable values. One of the parameters of this model formulation is the rate of development at a specific temperature within the optimal range. Ikemoto (2005) proposed a generalization to the Schoolfield et al. (1981) model which includes this reference temperature as a parameter.

1.2.2 Modelling Individual Variability

The models described in the section above all represent a population mean response, with no focus on individual variation in development rates. For utility in population dynamics modelling,

a phenology model should account for this variability across individuals. Two distinct categories of models used for variability are cohort-based and individual-based models. Cohort models are mainly based on observations of the overall distribution of a population across development stages, while individual-based models require repeated observations of development for each individual in a population (Chuine and Régnière, 2017). These individual observations can be summarized into population level observations, but the converse is not true. If samples are taken in the wild rather than in a laboratory setting, individual-level data and mortality data are nearly impossible to obtain. For these reasons many of the variability models that have been developed fall into the cohort category, despite the advantages of the individual-based approach (DeAngelis, 2018).

Some of the earliest models of individual variation treat hatch, larval stage progression and death stochastically, as states through which the population progresses. In Read and Ashford (1968), the proposed model deals with growth and survival simultaneously. The time X spent in stage S_i is represented as the minimum of the time taken to reach stage S_{i+1} and the time in stage S_i until death. Survival is modelled as a Poisson process, while the times in stage can follow any distribution with suitable support and assumptions. They proposed an Erlangian distribution, which is a Gamma distribution with integer shape parameter, while Kempton (1979) proposed a normal or inverse Gaussian distribution. There has also been work done using life-tables to model the number of individuals entering a stage (Dempster, 1961; Southwood and Jepson, 1962). Bellows Jr. and Birley (1981) built upon this model, including population density at each stage with individual development time following a positively skewed distribution.

The models described in the section thus far place distributional assumptions on development time, but do not provide any mechanistic representation of development. In contrast, Sharpe et al. (1977) use an assumption of normally distributed enzyme concentration to extend to a model with normally distributed instantaneous development rate. Overall development rate is represented as the constant development rate per unit of control enzyme, multiplied by the normally distributed concentration of enzymes. The formula for the former value comes from the biophysical model described by Sharpe and DeMichele (1977). Transforming the normal distribution for rates into times gives a positively skewed distribution of times, in accordance with observed data. The mean development rate was fitted using nonlinear regression, and the standard deviation was calculated from observed development rates.

1.2.3 Applications to the Spruce Budworm

Nonlinear Mean Model

Weber et al. (1999) fitted the model in Taylor (1981) to data from insects reared under laboratory conditions at constant temperatures. The insects were collected from several locations in Canada while the budworm populations were in outbreak, as far north as the Northwest Territories, and the model was fitted to each population separately and compared. Model parameters differed across populations from different locations, and large individual variation within populations was observed. A notable difference in this work from others like it is that development was not broken up by stage, but rather was measured from the beginning of the third larval stage all the way to pupation. Nonlinear least squares was used for parameter estimation.

Degree-Day Cohort Model

Dennis et al. (1986) proposed an extension of Osawa et al. (1983), a stochastic model of balsam fir bud phenology, to fit spruce budworm data observed in the wild. Osawa et al. (1983) assumed that development level (or "age") follows a normal distribution, with mean equal to the number of accumulated degree days and with variance proportional to the mean. Dennis et al. (1986) used a logistic distribution, since it can be similar in shape and the cumulative distribution function (CDF) has an analytic form. The model parameters include development thresholds for each stage and a variance multiplier. Development thresholds are the estimated ages in degree-days at which moults take place. Since it is difficult to determine the actual age of an insect in the field, the model only requires an observation of the developmental stage. The observations of age thresholds are therefore interval censored, so the probability that a randomly selected larva resides in a given stage is the difference of two normal CDFs. The parameters are estimated using maximum likelihood; the authors caution that multiple starting values should be used due to the potential complexity of the likelihood surface. Lysyk (1989) fits this model to several sets of spruce budworm data. The optimal lower threshold for calculating degree days was estimated by fitting the model for a range of threshold values, suggesting 8 degrees Celsius as the temperature that provided the best fit. This temperature threshold is different from that determined by Weber et al. (1999), and is also different from thresholds used in models in the following sections (Stedinger et al., 1985; Hudes and Shoemaker, 1988; Régnière, 1987). In Weber et al. (1999), the lower temperature threshold is a

model parameter, but the others estimate them either graphically or based on goodness-of-fit of the model parameters.

Nonlinear Cohort Models

Stedinger et al. (1985) also extend Osawa et al. (1983), using a cubic polynomial of degree days and incorporating inter-site variation instead of using only untransformed degree days. The mean model is a linear regression equation with an intercept term, a binary variable for tree type, a Julian date variable and the cubic polynomial of degree days. To account for variation in populations across sites, they use a Dirichlet model for the joint distribution of the proportions of the population in each stage for each site. Maximum likelihood estimates for each parameter were found using an iterative search technique, and the observed Fisher Information was used to estimate the standard errors of the parameters. Based on the likelihood ratio statistic, the model that included a cubic polynomial of degree days provided a better fit to the data than a linear function of degree days.

Hudes and Shoemaker (1988) proposed a generalized version of this model, using the logistic distribution proposed by Dennis et al. (1986) and allowing the mean model to be an arbitrary function of temperature. They ultimately used a regression equation similar to Stedinger et al. (1985) with the cubic polynomial of accumulated degree-days, a calendar date variable and a tree type variable and adding an indicator variable for geographic location.

Nonlinear Individual-Based Models

Régnière (1987) proposed a model with two components: one for the population median and another for individual variation around the median. The population median component uses the development rate model from Logan et al. (1976), with some slightly modified parameters. This median model is fitted separately for each stage of development. Individual variation is incorporated with a stage- and individual-specific multiplier of the mean development rate. The proposed CDF of this multiplier is a two-parameter logistic equation with median one. The median development component is fitted using the reciprocal of the observed times to moult, while the variation component is fitted to the ratio of each individual's development rate to the population median. The median and individual variation components of the model are fitted separately using nonlinear least squares and the components are both independent across development stages. This is stricter than the model in Hudes and Shoemaker (1988) who note that their model formulation allows for any correlation structure for

individual development across stages (Hudes, 1982). Since this model is individual-based, it requires sustained observations of the same individuals for the duration of their development.

An extension of this model is proposed in Régnière et al. (2012), wherein the mean and variation components of the model are combined in a likelihood framework. The common thread connecting the two models is the calculation of an individual's age; it is based on the same multiplier of the population median temperature-dependent development rate. A lognormal distribution is proposed to represent individual variation around the mean. The likelihood framework also accommodates censored observations and the estimation of parameters from population means rather than individual-level data, although more accurate estimates are obtained from individual-level data. The framework can also accommodate any development rate function; they include an example in which they fit the Schoolfield et al. (1981) model using their framework. A multiplicative population-level random effect is also included to account for differences across populations reared at different constant temperatures.

1.2.4 Parameter Estimation

All of the mean-rate nonlinear models described above are fitted using nonlinear least squares regression, many citing the iterative technique shown in Marquardt (1963). Often simplifications to models were proposed in order to ease the computational burden of estimating parameters, such as the model in Schoolfield et al. (1981). While many of the models incorporating individual variation were also fitted using nonlinear least squares, some authors used maximum likelihood to estimate parameters. Bayesian methods have successfully been used to fit phenology models for plants (Ashcroft et al., 2016; Fu et al., 2012; Iizumi et al., 2009; Thorsen and Höglind, 2010), microorganisms (Corkrey et al., 2012) and ectotherms (Feng et al., 2020).

1.3 Research Objective

The objective of this project is to implement a Bayesian hierarchical model for the individual-based, temperature-dependent development rates of the spruce budworm, and to compare the resulting parameter estimates across development stages and the colonies' geographic origins.

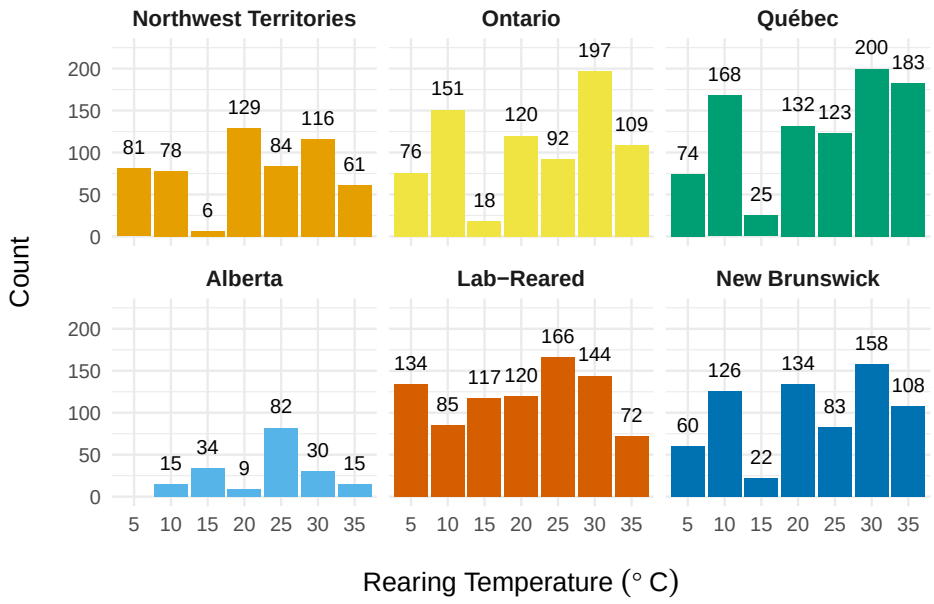
Chapter 2

Methods

2.1 Data Collection

The data for this project comes from a spruce budworm (*Choristoneura fumiferana*) rearing experiment described in Wardlaw et al. (unpublished manuscript). Samples of diapausing larvae were collected as described in Candau et al. (2019) from several locations across Canada: Inuvik, NWT; High Level, AB; Timmins, ON; Manic-5 weather station, QC; and Balmoral, NB (Figure 2.2). The sampling was performed in accordance with the methods in Perrault et al. (2021). Data was also collected from a colony obtained from the Insect Productions Services insectarium at the Great Lakes Forestry Centre in Sault Ste. Marie, ON; this colony had been reared under laboratory conditions for many generations. Upon emergence from diapause, individuals were collected and placed in individual containers containing artificial diet created to mimic the nutrition that the insects would consume in the wild. The development of each insect was observed daily; a moult was reported once a larva had shed its head capsule. In order to observe development as a function of temperature, each colony was divided into seven sub-populations. Each sub-population was reared at a different constant temperature, at seven evenly spaced temperatures from $5 - 35°$C. Since larval development of Canadian spruce budworm populations occurs mostly in the spring and summer, the "optimal" temperature range was determined to be $15 - 25°$C. Due to the potential for high mortality rates at temperatures outside an optimal range, an extra rearing step was taken for populations at extreme temperatures. For each stage, the times to moult at the extreme temperatures were estimated using BioSIM (Régnière et al., 2014). The insects were held at those temperatures for half this time, and

Figure 2.1: Sample sizes at each treatment level

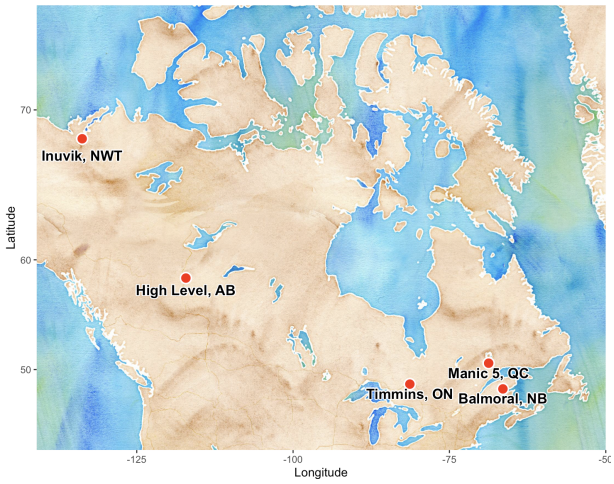


then moved to 20°C for the remainder of each stage. Each level of the rearing experiment contained 250 individuals, resulting in a total of 10,500 individuals. Since the main objective of the project was to observe development rates in the larval instar stages, survival was not considered in the modelling process. Thus, the data used to fit the model consisted only of individuals who survived the full rearing process from the first larval instar to pupation; the counts of surviving individuals at each treatment level are displayed in Figure 2.1. The overwintering stage was also outside the scope of this project. Only the first new generation from the wild populations was used to fit the data, to control for any generational effect of laboratory rearing.

2.2 Model Implementation

There are several ways to implement Bayesian hierarchical models in R. Commonly-used packages for this purpose are JAGS, WinBUGS, Nimble, Stan and TMB. Due to its computational efficiency, Hamiltonian Monte Carlo (HMC) was preferred over Gibbs Sampling. For this reason, TMB and Stan were preferred over the other solvers. The model was written in TMB; the R package `tmbstan` (Monnahan and Kristensen, 2018) was used to apply Stan's No-U-Turn Sampler (NUTS), a variant of HMC, to the model.

Figure 2.2: Collection sites of wild spruce budworm populations across Canada



2.2.1 Hamiltonian Monte Carlo

HMC is a Markov Chain Monte Carlo (MCMC) algorithm that simulates the trajectory of a particle moving under Hamiltonian dynamics to explore a desired target distribution. Using an auxiliary random variable to represent momentum, the algorithm uses the gradient of the target surface along with the momentum in each dimension to explore the surface and to propose new sample points. In the traditional HMC algorithm, a number of leapfrog steps L and a step size ϵ are specified in advance. Each simulated trajectory runs for L discrete steps along the surface for time ϵ, and the final point is either accepted or rejected using a Metropolis acceptance criteria. The No-U-Turn Sampler (NUTS) (Hoffman and Gelman, 2014) is an extension of the HMC algorithm that automatically optimizes step size and dynamically adjusts the number of leapfrog iterations required for a sample. Instead of specifying the number of steps L in advance, the algorithm runs until it detects that the trajectory is moving back towards the point of origin.

2.2.2 Diagnostics

Ideally, these solvers return posterior samples that are as close to independent as possible, with enough samples that they are sufficiently representative of the full distribution. While convergence to the true posterior distribution cannot be fully confirmed, there are diagnostics that can be used to assess convergence and thus to inform the user whether their samples are to be trusted. The three main numerical diagnostics in Stan are the Gelman-Rubin statistic ($\hat{R}$), the effective sample

size (ESS), and the number of divergent transitions in the sample.

Introduced by Gelman et al. (1992), $\hat{R}$ is a measurement of how well distinct MCMC chains have mixed, and thus whether the sampler has achieved adequate convergence. The statistic is based on the ratio of the between-chain variance to the within-chain variance. The estimator is written as

$$\hat{R} = \sqrt{\frac{\frac{N-1}{N}W + \frac{1}{N}B}{W}}$$

where N is the chain length, W is the within-chain variance and B is the between-chain variance. If the between-chain variance is large in comparison to the within-chain variance, the chains have not mixed adequately. Though this diagnostic is widely used, it was shown by Vehtari et al. (2021) to be seriously flawed. They propose using the same statistic, but applied to the rank-normalized values of the posterior draws. Using ranks instead of raw values ensures that the statistic can still diagnose biases in posterior samples with infinite variance. They also recommend calculating the $\hat{R}$ statistic for the absolute deviations from the across-chain median of the rank-normalized values. If one chain is not properly exploring the tails of the posterior distribution, its variance will be smaller than that of the other chains. While the $\hat{R}$ statistic taken on the untransformed posterior draws would not detect this difference, transforming the values into folded deviations from the median makes the mean of that chain smaller than the rest, thus enabling the diagnostic to detect non-convergence. Vehtari et al. (2021) propose taking the maximum of the two $\hat{R}$ statistics described (rank-normalized and absolute deviations from median), and recommend accepting the samples only if the value is smaller than 1.01.

Vehtari et al. (2021) propose a similar transformation to calculate the ESS. For estimation of a parameter mean, the posterior sample contains as much information as an independent sample that is the size of the ESS. The ESS is a function of the autocorrelations in the Markov chains. It is written as

$$N_{eff} = \frac{N}{1 + 2\sum_{t=1}^{\infty} \rho_t}$$

where N is the chain length and ρ_t is the autocorrelation at lag t. By rank-normalizing the posterior sample, the normality assumptions of both statistics are satisfied; they are also invariant to monotone transformations. The traditional ESS used is referred to as the "bulk-ESS", as it is only representative of the effective sample size for the bulk, or mean, of the distribution. This means that a separate statistic must be calculated to evaluate the ESS in the tails of the distribution. For this calculation,

the quantiles of interest from the posterior distribution are calculated using the empirical CDF. The posterior samples are transformed using an indicator function, and the ESS is calculated on the resulting sample. Vehtari et al. (2021) define the "tail-ESS" as the minimum ESS of the 5th and 95th quantiles of the posterior sample. The overall ESS is the minimum of the tail-ESS and the bulk-ESS.

Divergent transitions occur when a particle's estimated trajectory is determined to have departed significantly from its true trajectory; this is determined by the total energy in the system. The energy of the system is a sum of potential and kinetic energy; the former is a function of the particle's position, and the latter is a function of its momentum. Since total energy should be conserved throughout the system, a transition is deemed divergent when the energy calculated from the end of the simulated trajectory deviates significantly from the energy at the beginning of the trajectory. The appearance of divergent transitions indicates that there are areas of the posterior distribution that the solver cannot explore properly, often due to high curvature. Sometimes this is due to an overly large step; if there is high curvature, a large step size can send the particle off of a ridge on the target surface. In this case, decreasing the step size via an acceptance probability hyperparameter can help the solver to explore those regions more carefully. However, divergent transitions can also indicate issues with model implementation. In hierarchical models, a phenomenon known as Neal's Funnel (Neal, 2003) can occur when the joint density of the two variables in the hierarchy is "squeezed" by the lower parameter. For instance, if the variance of a normally distributed parameter is itself a parameter, small values of the variance parameter can lead to very concentrated values of the draws from the higher-level distribution. This can create a steep surface in the parameter space, potentially giving rise to divergent transitions. A solution to this issue is to de-centralize the hierarchical variables in the model; that is, sample each parameter separately from a simpler distribution with fixed parameters, and transform the variables after they are sampled to obtain the values required for the model. In the previously stated example, the higher-level variable could be sampled from a standard Normal distribution, multiplied by the sampled lower-level standard deviation parameter and added to the desired mean. This effectively flattens the surface of the target distribution, making it easier to explore (Stan Development Team, 2020).

2.3 Model Development

The implementation of the model is based on the likelihood framework presented in Régnière et al. (2012). This framework is, in turn, based upon a series of papers from Régnière, culminating in the nonlinear, individual-based development rate model in Régnière (1987).

2.3.1 Original Two-Component Development Rate Model

The spruce budworm development model used in BioSIM (Régnière et al., 2014) is outlined in Régnière (1987). This individual-level model has two components: one for the population median and another for individual variation around the median. In the model, the developmental "age" of an individual i in stage j is represented as

$$a_{ij} = \delta_{ij} \int_0^t r(T_{t'})\, dt' \tag{2.1}$$

where $r_j(T_t)$ represents the median instantaneous development rate at the temperature T experienced at time t, and δ_{ij} is a multiplier for an individual's development rate from the median of individuals in stage j. An individual's age in a given development stage is the proportion of the stage it has completed; 0 at the beginning of the stage and 1 at the end. The units of the development rate are therefore time^{-1}. Taking the sum of the instantaneous development rates at the temperatures experienced by an individual, weighted by the amount of time the individual spends at each temperature, gives its developmental age.

The nonlinear equation in Logan et al. (1976) with simplified parameters is used to represent the median development rate

$$r_j(T_t) = \begin{cases} \beta_1 \left[\frac{1}{1+e^{\beta_2 - \beta_3 \tau}} - e^{(\tau-1)/\beta_4} \right] & \text{if } T_b \leq T \leq T_m, \text{ where } \tau = \frac{T - T_b}{T_m - T_b} \\ 0 \text{ otherwise} \end{cases} \tag{2.2}$$

where T_b is the lower temperature threshold, and T_m is the upper threshold. Individual variation around the median is represented by the random variable δ_{ij}. In the model framework, the variable δ_{ij} can be assigned any distribution. In Régnière (1987), the cumulative distribution function used to model individual variation is

$$F_\delta(d) = [(2^q - 1)e^{k(1-d)} + 1]^{-1/q} \tag{2.3}$$

where q and k are shape and scale parameters. Setting $F_\delta(d)$ to 0.5 and solving for δ shows that the median of δ is one, which ensures that the median development rate of the population can be accurately represented by r.

There are a few significant assumptions underlying this model. The first is that the effect of temperature on development rate is strictly additive. That is, an insect's development rate at a given temperature is independent of the temperatures previously experienced. This is likely an oversimplification of reality, but weakening this assumption is difficult when dealing with strictly constant temperature data. A variable temperature regime is required to detect differences in instantaneous development rates caused by previous temperatures experienced. Nevertheless, nonlinear rate summation is more flexible than degree day modelling; it relaxes the assumption that the relationship between development and temperature is linear.

Another assumption in this model is that an individual's variation from the population's median development rate is independent across development stages. Since the median development and individual variability components are fit separately over stages, there is nothing connecting an individual's variation multiplier δ from one stage to the next. Furthermore, implicit in this model formulation is the fact that within a stage, an individual's development at a given temperature is proportional to its development at any other temperature. That is, if individual i in stage j has multiplier $\delta_{ij} = 1.5$, then that individual is expected to develop at a rate 1.5 times faster than the population median at any temperature.

The reality of measuring development rates at extreme temperatures under a constant temperature regime necessitates another assumption under the constraints of this model. Due to the temperature spiking method used to obtain this data, there is an issue of non-identifiability that arises in estimating development rates. Namely, an individual spends a fraction of each life stage in the extreme temperature at which inference is to be made, and the other fraction at a temperature that is more suitable for development. Since the only value observed is total time in stage, the true proportions of development in each temperature are not observed and must instead be imputed.

Régnière (1987) provides a simple fix for this problem. He suggests using the estimate of the median development rate at the optimal temperature as the imputed value to estimate the devel-

opment rate at the extreme temperature. Since a stage is considered complete when an individual reaches age 1, the following equation represents the breakdown of development occurring in a stage:

$$1 = r_e t_e + r_o t_o \tag{2.4}$$

Here, r_e is the development rate at the extreme temperature (the value of interest), r_o is the development rate at the optimal temperature (the imputed value), and t_e and t_o represent the days spent at extreme and optimal temperatures, respectively. Rearranging this gives

$$r_e = \frac{1 - r_o t_o}{t_e} \tag{2.5}$$

The value of r_e can be estimated by plugging in a suitable estimate for r_o. In this case, it is necessary to assume that the population of individuals reared at the extreme temperature have the same mean development rate at the optimal temperature as the population that was fully reared at the optimal temperature. This is not likely to be exactly true, but the estimate r_o is the best estimate of true value of the extreme temperature population's median development rate at the optimal temperature.

In Régnière (1987), the two model components are fitted separately using nonlinear least squares regression. Each of these fits is performed for each stage independently. The mean model for a given stage is fitted to the reciprocal values of the observed times in the stage. If there are unbalanced observations across different temperature treatments, the regression can be weighted by the number of individuals in each treatment. The variability component of the model for a given stage is fitted using the ratio of the observed development rates to the median development rate in a treatment. These ratios are pooled across temperature treatments, and the parameters of the cumulative distribution function of δ are estimated using the empirical quantiles of the ratios.

Estimating Development Rates from Transfer Treatment Observations

The method used in Régnière (1987) to impute r_e relies on the assumption that each individual completes the same proportion r_o of the stage per day at the optimal temperature. However, according to previous model assumptions, r_o for a given individual should be proportional to r_e for the same individual. If Equation 2.5 is re-written for individual i where r_e^* and r_o^* are the population median

development rates at the extreme and optimal temperatures, it becomes

$$1 = \delta_i(r_e^* t_{e,i} + r_o^* t_{o,i}) \tag{2.6}$$

Thus, a better estimate of r_e requires the imputation of δ_i as well as r_o.

A more suitable estimate for r_e^* can be obtained using the method of moments. If we assume that the sample mean of δ_i is a suitable estimate for its expected value, Equation 2.6 can be rearranged as

$$\delta_i = \frac{1}{r_e^* t_{e,i} + r_o^* t_{o,i}} \tag{2.7}$$

and so

$$E(\delta) \approx \frac{1}{N} \sum_{i=1}^{N} \frac{1}{r_e^* t_{e,i} + r_o^* t_{o,i}} \tag{2.8}$$

where N is the number of insects in the extreme treatment and r_o^* is estimated from Régnière's suggested method. The parameters of F_δ are universal for a given stage, so their estimates are pooled across all of the populations reared at optimal temperatures. Estimates of those parameters can lead to a reasonable estimate of $E(\delta)$, so solving Equation 2.8 for r_e^* gives an estimate of the development rate at the extreme temperature.

2.3.2 Likelihood Framework

In Régnière et al. (2012), the multiplicative structure for individual variation described in the section above is used to obtain a likelihood framework that can be used in conjunction with any development rate model and distribution of individual variation. Since observations occur at discrete intervals, the exact time a moult occurs is unknown; all observations are interval censored. Given a vector of model parameters $\boldsymbol{\theta}$ and an observed moult at time t, the likelihood of the observation is

$$L(\boldsymbol{\theta}; t) = P(t - \Delta t < X < t) \tag{2.9}$$

where X is a random variable representing the time the moult actually took place and Δt is the time since the previous observation. The distribution of X can be derived from Equation 2.1. For

an individual reared under a constant regime at temperature T, this equation can be re-written as

$$1 = \delta r(T, \boldsymbol{\theta}) X \tag{2.10}$$

since the age of an individual at the completion of a stage is 1. Rearranging gives $X = \frac{1}{\delta r(T,\boldsymbol{\theta})}$, and thus the probability can be written as

$$L(\boldsymbol{\theta}; t) = P\left(t - \Delta t < \frac{1}{\delta r(T, \boldsymbol{\theta})} < t\right) = P\left(\frac{1}{r(T, \boldsymbol{\theta})t} < \delta < \frac{1}{r(T, \boldsymbol{\theta})(t - \Delta t)}\right) \tag{2.11}$$

Since the distribution of δ is known, 2.11 can be expressed as

$$L(\boldsymbol{\theta}; t) = F_\delta\left(\frac{1}{r(T, \boldsymbol{\theta})(t - \Delta t)}\right) - F_\delta\left(\frac{1}{r(T, \boldsymbol{\theta})t}\right) \tag{2.12}$$

This likelihood holds for any development rate function $r(T)$ and any distribution of δ, F_δ.

This probability framework can also accommodate extreme temperature transfer data. In the transfer treatments, individuals begin each stage at the extreme temperature for a pre-determined amount of time, and are transferred to a more optimal temperature for the remainder of the stage. Thus, Equation 2.10 becomes

$$1 = \delta \left[r(T_o, \boldsymbol{\theta}) X + r(T_e, \boldsymbol{\theta}) t_e\right] \tag{2.13}$$

where T_e and T_o represent the extreme and optimal treatment temperatures respectively, and t_e is the time spent at the extreme temperature. Equation 2.12 then becomes

$$L(\boldsymbol{\theta}; t_o, t_e) = F_\delta\left(\frac{1}{t_e r(T_e, \boldsymbol{\theta}) + (t_o - \Delta t) r(T_o, \boldsymbol{\theta})}\right) - F_{\delta_j}\left(\frac{1}{t_e r(T_e, \boldsymbol{\theta}) + t_o r(T_o, \boldsymbol{\theta})}\right) \tag{2.14}$$

where t_o is the observed time spent at the optimal temperature. In some instances, the moult occurs during the extreme temperature period. In this case, the strictly constant temperature formulation can be used.

A lognormal distribution for δ is proposed, since it is typically positively skewed, it ensures $\delta > 0$, and its multiplicative inverse is also lognormal with parameters of the same magnitude. This means that there should be no difference between modelling rate and modelling time. Régnière et al. (2012) also include a multiplicative, normal random effect υ to account for lack of fit between the theoretical development rate and the treatment population mean. While this model formulation maintains many

of the assumptions of the model in Régnière (1987), it provides a more elegant method for dealing with temperature transfers, and it can be parameterized using maximum likelihood estimation.

2.3.3 Development Rate Model

The Schoolfield et al. (1981) model was chosen for the Bayesian implementation of the Régnière et al. (2012) likelihood framework. The development rate equation in this model is

$$r(T_K) = \frac{\rho_{25}\frac{T_K}{298}\exp\left[\frac{H_A}{R}\left(\frac{1}{298} - \frac{1}{T_K}\right)\right]}{1 + \exp\left[\frac{H_L}{R}\left(\frac{1}{T_L} - \frac{1}{T_K}\right)\right] + \exp\left[\frac{H_H}{R}\left(\frac{1}{T_H} - \frac{1}{T_K}\right)\right]} \tag{2.15}$$

where $R = 1.987\text{kcal} \cdot \text{K}^{-1} \cdot \text{mol}^{-1}$ is the universal gas constant and T_K is the input temperature in degrees Kelvin. All temperatures in this model are represented in degrees Kelvin. The formula represents the enzyme-catalyzed reaction governing growth rate. It is based on the assumption that these rate-controlling enzymes can either be active, or inactive due to low temperature or high temperature inactivation; it is assumed that these states are reversible. The reciprocal of the denominator represents the fraction of enzymes in the active state.

The parameters as described as follows:

ρ_{25} : development rate at 25°C (298 K, estimated as an optimal development temperature)

T_L : temperature at which the enzyme population is half active and half low-temperature inactive

T_H : temperature at which the enzyme population is half active and half high-temperature inactive

H_L : change in enthalpy associated with low-temperature enzyme inactivation (cal mol^{-1})

H_H : change in enthalpy associated with high temperature enzyme inactivation (cal mol^{-1})

H_A : enthalpy of activation of the reaction catalyzed by the enzyme (cal mol^{-1})

The enthalpy of activation is analogous the energy required for a reaction to take place; the larger this value, the slower the reaction. Depending on its sign, change in enthalpy represents either a gain or loss of energy in a system. The choice of the reference temperature of 298K (or 25°C) was carried forward from Schoolfield et al. (1981); enzyme inactivation due to extreme temperatures is expected to be minimal at this reference temperature.

This model provides some important advantages over other options, specifically within a Bayesian

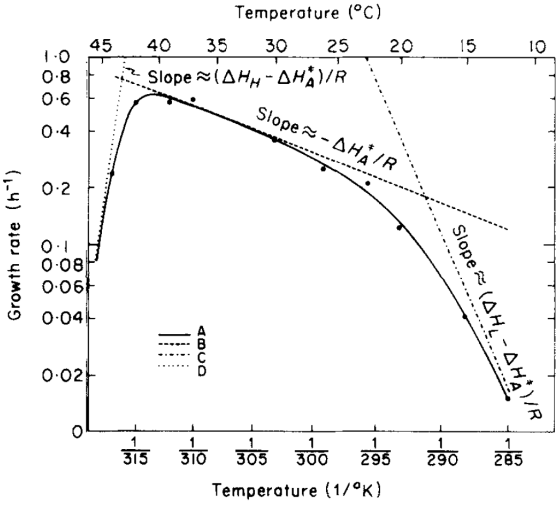
framework. The first is the absence of temperature thresholds; while many of the development models in the literature include temperature thresholds, there is little evidence that they are crucial to modelling development (Chuine and Régnière, 2017). While there is evidence of increased mortality at extreme temperatures, mortality can be seen as a separate process from development. Under the structure of the Régnière et al. (2012) model, individuals vary around a single population mean development rate curve. Since the temperature thresholds are features of this mean curve, each individual would be assumed to have the same threshold. A fully continuous development rate curve that approaches zero asymptotically at the extreme temperatures provides more flexibility, since development can effectively drop to zero for certain individuals while allowing others to continue development. Discontinuity is another disadvantage of certain implementations of temperature thresholds, especially when dealing with gradient-based solvers such as Stan. Not only does it make the surface non-differentiable in certain areas, but the log-likelihood becomes unstable. For instance, if a sampled lower temperature threshold is greater than a point where development occurs, the log-likelihood becomes negative infinity. This makes it difficult to specify a prior for these threshold parameters.

The other advantage of the Schoolfield et al. (1981) model is the interpretability of the parameters. Most parameters are rates and temperatures, for which estimates of central tendency and spread are intuitive. The enthalpy parameters, H_L, H_A and H_H, are not easily estimated but Sharpe and DeMichele (1977) provide estimates with starting values based on averages over more than 50 enzyme-substrate systems. The scale of the prior distributions can be estimated based on the magnitude of the values. Graphical interpretations of these parameters are also explored in Schoolfield et al. (1981); they can be used to approximate the slopes of three different sections of the curve when plotting the reciprocal of temperature in degrees Kelvin on the x-axis and the log development rate on the y-axis, as depicted in Figure 2.2. On a plot of temperature and development rate, they control the slope of the curve at the points T_L, 298 K and T_H.

2.4 Hierarchical Model

The remaining model parameters represent individual variation around the mean development curve, a random effect representing for lack-of-fit between the theoretical development rate response and the treatment mean, and a variance parameter for the random effect. Each temperature treatment

Figure 2.3: Graphical interpretations of H_L, H_A and H_H, from Schoolfield et al. 1981



in the dataset is modelled as a separate treatment, since each of those subpopulations are reared separately. Therefore, each level of the random effect for the deviation of the treatment mean from the estimated development rate represents a different rearing temperature.

These remaining model parameters are defined as:

σ_ϵ : standard deviation of the random variable $\epsilon = \log(\delta)$

υ : multiplicative lack-of-fit between theoretical response and treatment mean (random effect)

σ_υ : standard deviation of υ

The full log-likelihood function of the parameter set $\boldsymbol{\theta}$ for one population level (i.e. one stage of one colony) is

$$ll(\boldsymbol{\theta}) = \sum_j \left\{ \log\left[f_\upsilon(\upsilon_j)\right] + \sum_i \log\left[F_\delta\left(\frac{1}{\upsilon_j\left[t_{e_{ij}} r(T_{e_j}, \boldsymbol{\theta}) + (t_{o_{ij}} - \Delta t_{ij}) r(T_{o_j}, \boldsymbol{\theta})\right]} \right) - F_\delta\left(\frac{1}{\upsilon_j\left[t_{e_{ij}} r(T_{e_j}, \boldsymbol{\theta}) + t_{o_{ij}} r(T_{o_j}, \boldsymbol{\theta})\right]} \right) \right] \right\} \tag{2.16}$$

where f_υ is the normal density function with mean 1 and variance σ_υ^2, and F_δ is the log-normal cumulative distribution function where the normal random variable $\epsilon = \log(\delta)$ has mean $-\frac{1}{2}\sigma_\epsilon^2$ and variance σ_ϵ^2. The latter distribution implies that δ is log-normally distributed with mean 1. The

rate equation r is represented by Equation 2.15. The i and j subscripts represent individuals and treatments, respectively. In this model, the treatments correspond to different rearing temperatures. The value Δt_{ij} represents the amount of time between the each observed moult and the previous observation.

2.4.1 Additional Structure

Upon examining the posterior samples for each population level, it was evident that there was a strong quadratic relationship between stage and ρ_{25}. To fit this structure, the ρ_{25} parameter was replaced by a quadratic function of the discrete development stage. Using stages as integer values provides a simplified discrete approximation to where the population actually lands in the life cycle. This discrete approximation could eventually be replaced by a continuous proxy for life cycle progress, such as insect weight. To model the quadratic structure, the coordinates of the vertex and the ratio, ϕ, of the development rate at the vertex to the development rate at the final development stage were used as parameters in the quadratic equation. In vertex form, the equation for ρ_{25} is

$$\rho_{25} = a(s - x_v)^2 + y_v \tag{2.17}$$

where s is stage ($s = 0$ for L2, $s = 1$ for L3, etc.), x_v and y_v are the vertex coordinates, and a is an amplitude parameter. Letting ϕ be the ratio mentioned above, the curve should pass through the point $(4, \phi y_v)$, so

$$\phi y_v = a(x_v - 4)^2 + y_v \tag{2.18}$$

Solving for a gives

$$a = \frac{y_v}{(x_v - 4)^2}(\phi - 1) \tag{2.19}$$

and thus the quadratic representation of ρ_{25} is

$$q(s) = \frac{y_v}{(x_v - 4)^2}(\phi - 1)(s - x_v)^2 + y_v \tag{2.20}$$

To add flexibility to this structure, a multiplicative lack-of-fit parameter was included as a random

effect. Where ρ_{25} appeared in the original model, it is replaced by

$$\rho_{25} = q(s)e^{\alpha} \tag{2.21}$$

where $\alpha \sim N(\mu = -\frac{\sigma_\alpha^2}{2}, \sigma = \sigma_\alpha)$ to ensure that $E(e^\alpha) = 1$.

The remaining model parameters are treated as independent across stages, maintaining the flexibility for each parameter that is attained by running each level separately, while combining the stages through the quadratic function for the ρ_{25} parameter.

2.4.2 Choice of Priors

Original Model

Most of the priors specified for the model are centred at the values published in Régnière et al. (2012). The priors are defined as follows:

$$\begin{aligned}
\rho_{25} &\sim \text{Beta}(\alpha = 8, \beta = 8) \\
T_L &\sim \text{Normal}(\mu = 283.9, \sigma = 2) \\
T_H &\sim \text{Normal}(\mu = 306.2, \sigma = 2) \\
H_L &\sim \text{Normal}(\mu = -59,700, \sigma = 1,000) \\
H_A &\sim \text{Normal}(\mu = 11,400, \sigma = 5,000) \\
H_H &\sim \text{Normal}(\mu = 100,000, \sigma = 20,000) \\
\log(\sigma_\epsilon) &\sim \text{Normal}(\mu = -0.75, \sigma = 0.075) \\
\log(\sigma_\upsilon) &\sim \text{Normal}(\mu = -2.5, \sigma = 0.1) \\
\upsilon &\sim \text{Normal}(\mu = 1, \sigma = \sigma_\upsilon)
\end{aligned}$$

In the scaled, de-centralized model implementation, certain priors are altered for ease of computation. The affected priors are as follows:

$$H_L \cdot 0.001 \sim \text{Normal}(\mu = -5.97, \sigma = 0.1)$$
$$H_A \cdot 0.001 \sim \text{Normal}(\mu = 1.14, \sigma = 0.5)$$
$$H_H \cdot 0.001 \sim \text{Normal}(\mu = 10, \sigma = 2)$$
$$\eta \sim \text{Normal}(\mu = 0, \sigma = 1)$$
$$\upsilon = \eta\sigma_\upsilon + 1$$

Since ρ_{25} represents a development rate and completion of a stage is not expected to take less than one day, the $[0, 1]$ support of the beta distribution makes it an appropriate prior for this variable. The variances of the temperature parameters were determined based on the typical spread of temperature values experienced by the spruce budworm, while the spread of the enthalpy parameters was based simply upon the magnitude of the published parameter values. The signs of these parameters are known; H_A and H_H should be positive while H_L should be negative. The chosen standard deviation makes sign changes for these priors rare. Since the variance parameters must be positive, the priors were set for the log-transformations of these variables to ensure that they had the proper support. The σ_ϵ parameter represents the deviation of individuals from the mean development rate curve, while σ_υ represents the deviation of a treatment mean from the expected mean. Variance of individuals around the mean is expected to be larger than that of treatments around the expected value. Since the objective function contains the log-transformed difference of evaluations of the normal CDF with standard deviation σ_ϵ, a small variance can lead to minute values of the difference and thus values of the objective function that are inflated in magnitude. These values can cause issues with for the solver, as the gradient of the likelihood surface can become too steep; suggesting larger values of this variance parameter mitigates this issue. The objective function to be minimized by the solver is the negative of the sum of the log likelihood and the log density of the prior draws.

Structured Model

The prior distributions on the parameters in the structured model remain the same as in the original model, with the exception of the ρ_{25} parameter, and the addition of extra parameters. The ρ_{25} parameter is replaced by the three parameters in the quadratic equation, and the random effect and

its standard deviation are also added as parameters. They are specified as follows:

$$
\begin{aligned}
x_v &\sim \text{Normal}(\mu = 2, \sigma = 0.1) \\
y_v &\sim \text{Beta}(\alpha = 5, \beta = 6) \\
\phi &\sim \text{Beta}(\alpha = 5, \beta = 5) \\
\log(\sigma_\alpha) &\sim \text{Normal}(\mu = -2, \sigma = 0.1) \\
\alpha &\sim \text{Normal}(\mu = -\frac{\sigma_\alpha^2}{2}, \sigma = \sigma_\alpha)
\end{aligned}
$$

and for the de-centralized implementation,

$$
\begin{aligned}
\log(\sigma_\alpha) &\sim \text{Normal}(\mu = -2, \sigma = 0.1) \\
\psi &\sim \text{Normal}(\mu = 0, \sigma = 1) \\
\alpha &= \psi\sigma_\alpha - \frac{\sigma_\alpha^2}{2}
\end{aligned}
$$

Since the L4 ($s = 2$) development stage is known to be the shortest, the prior for x_v was centred there. The initial prior on x_v was wider, with a standard deviation of one, but this led to samples with infinite log likelihoods. The width of the prior was decreased iteratively until the solver returned acceptable samples. Similar to the prior for ρ_{25} in the unstructured model, the support for y_v should also be in [0,1], so a Beta prior was chosen. Since it is meant to represent a fraction of y_v, a Beta prior was also chosen for ϕ.

2.4.3 Input Data

For each observation in a given stage and colony, the solver takes as input: extreme treatment temperature, optimal treatment temperature, observed time in extreme treatment, observed time in optimal treatment, previous observed time, number of observations and block index. The previous observed time is a value calculated from the other observations, under the assumption that the observation period is one day. For observations from optimal treatment temperatures, it is calculated by subtracting one day from the observed time in optimal treatment. For these observations, the observed time in extreme treatment is zero. For observations from extreme treatment temperatures,

there are two cases. The first is the typical case, where the moult occurs after the time at extreme temperature, during the time at optimal temperature. In this case the previous observed time is the time at optimal temperature, minus one day. The second case is that a moult occurs during the time at extreme temperature, before spending any time at optimal temperature. These observations are treated like observations for individuals in optimal temperature treatments. The optimal temperature entry is set to the extreme temperature value and extreme temperature entry is set to zero. The previous observed time is recorded as the observed time in treatment, minus one day. Due to the impossibility of an instantaneous moult, previous observed time values equal to zero are increased to 0.1, both for numerical convenience and realism.

Chapter 3

Results

3.1 Diagnostics

The model was run separately for each combination of stage and colony. Four chains were run for each level, with each chain running for 10,000 iterations including a 500-iteration warmup period. For each run, the maximum $\hat{R}$, minimum ESS and number of divergent transitions were calculated and returned. The values of the first two diagnostics for each of the levels were within acceptable ranges, with all values of $\hat{R}$ less than 1.01 and all values of the ESS greater than 10,000.

3.1.1 Divergent Transitions and σ_ϵ

After decreasing the solver's step size and de-centralizing the model's hierarchical parameters, the solver returned one divergent transition when fitting the model to the L4 stage of the laboratory-reared colony. A possible reason for this divergent transition is the difference of CDFs taken in the objective function, which is included to represent the interval censoring in the data. The distribution in question is that of the log-transformed multipliers representing individual variability. For an individual with very slow development relative to the population mean, the difference is taken in the lower tail of the CDF. The further into the tail the point lies, the flatter the CDF is at that point. Thus, the difference in the CDF evaluated at that point and a point nearby can be arbitrarily close to zero. As the log-likelihood is a function of the log of this difference, extreme points can cause its magnitude to increase dramatically. This suggests that the variance parameter of the log-transformed individual variability could be the source of divergent transitions, which can occur when

such steep features appear on the likelihood surface.

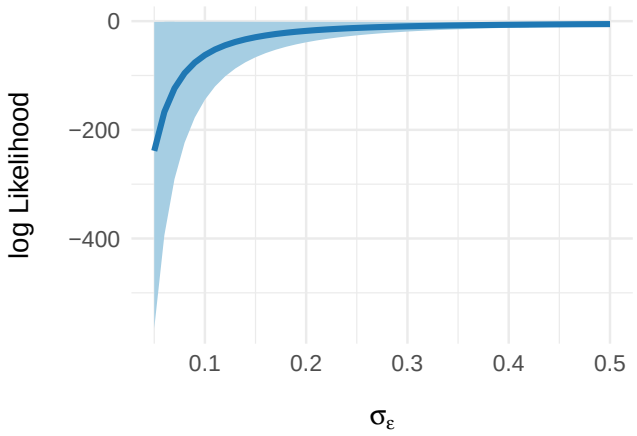


Figure 3.1: Calculation of likelihood for a point from the lab-reared colony at 25°C, for which stage L4 took 7 days. Light blue ribbon represents range of estimates over 100 samples from the prior and dark blue line represents median at each value of σ_ϵ.

Figure 3.1 illustrates how σ_ϵ can affect the likelihood of extreme observations. Increasing σ_ϵ should have the effect of pulling every observation out of the tails, into the bulk of the distribution. The marginal effect of σ_ϵ on the objective function decreases dramatically as σ_ϵ increases. Thus, a potential solution for the issue of divergent transitions in this model is to tighten the prior of $\log(\sigma_\epsilon)$ around a larger value. Setting the prior to $\log(\sigma_\epsilon) \sim \mathrm{N}(-0.7, 0.075)$ eliminated divergent transitions for the problematic population, and did not cause any issues when fitting the model to the rest of the populations.

3.2 Posterior Predictive Checks

Figure 3.2 shows estimates from the model fits for each colony and stage with data overlaid. The black points represent development rate estimates from the observed data. For non-transfer-treatment data, the development rates are estimated by subtracting 0.5 days from the observed number of days to moult, and taking the reciprocal. The subtraction is a rough adjustment for using interval censored data. For the transfer treatment data, the estimates are obtained using the method of moments as described at the end of Section 2.3.1, with negative development rate estimates set to zero. For extreme temperature treatments, the input values for days at optimal temperature are also adjusted by subtracting 0.5 days. Since these observations of development rate are estimated,

some discrepancies between these values and the model estimates can be expected. The dark blue lines in the figure represent the median development rate estimates from the posterior sample at each temperature. The transparent dark blue ribbons around these lines represent the interquartile range of estimates and the 5th-95th quantile intervals of the estimates.

The estimated development rate values at each of the rearing temperatures were transformed by the v random effect multipliers, and interpolation splines were used to connect the points. This was done using the `interpSpline` function in R's `splines` package (R Core Team, 2020). These interpolation splines were used as an alternative to the estimated curves themselves to better represent the fit of the curves to the data. If the v parameter is left out of the plotting completely, the graphic does not accurately represent the model fit, since the measured lack of fit to the data is not taken into account.

The innermost light blue ribbons represent the 2.5th and 97.5th quantiles from the distribution of individual variation around the median development rate estimates (the dark blue line). The lighter shades of the light blue ribbons are analogous to the dark blue ribbons, representing the 25th-75th and 5th-95th quantiles of the posterior sample. Using each posterior sample's respective value of σ_ϵ, the quantiles of the distribution of individual variation were calculated and multiplied by each value of the spline to obtain the upper and lower boundaries of each ribbon. We would expect most of the points to fall within the outermost light blue ribbons and that the dark blue ribbons pass through the centres of the point clusters at each temperature.

For the majority of levels most points lie within the outermost intervals, and most curves pass through the centres of the observed values. However, there are many instances where the median development rate curves and ribbons miss the data points completely. For instance, the L2 estimates for the Northwest Territories, Québec and Ontario colonies dip below all of the observations at 30 and 35°C. This also occurs for many populations at the 5°C observations. In the New Brunswick fits, the estimated observed development at 35°C is consistently zero, but the model estimates never reach those observations.

The thickness of the dark blue intervals show that the estimated development rates in the L2 stage are more variable than the estimates in the other stages. The non-extreme temperature estimates (15-25°C) show very little variability, but the estimates at both low and high extremes appear to be quite variable. For most other stages, there is little variability in development rate estimates across the full rearing temperature spectrum. Since most of the discrepancies occur at the transfer

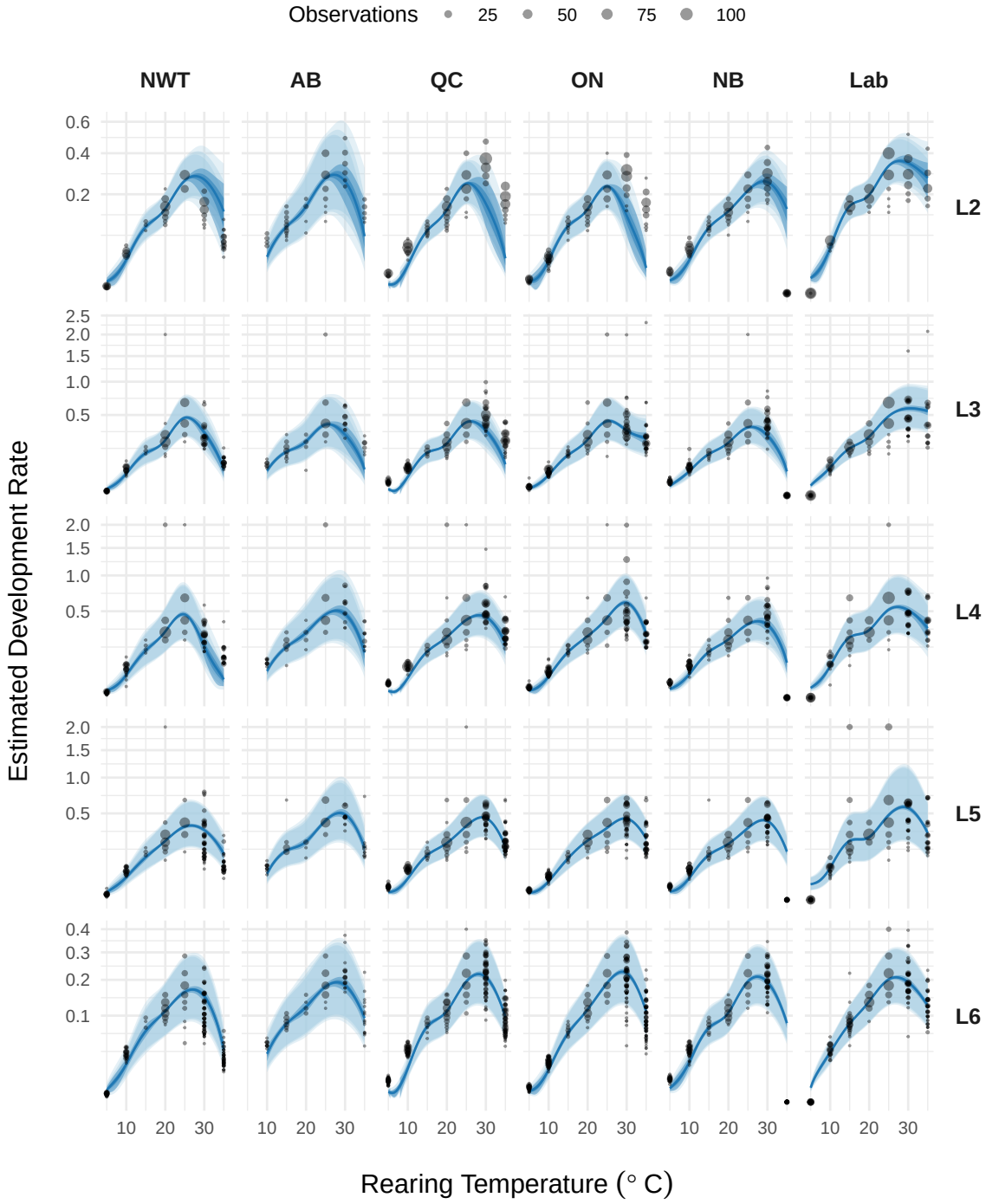


Figure 3.2: Posterior predictive checks. Dark blue lines represent median development rate estimates, shades of dark blue ribbons represent interquartile range and 5th-95th quantiles. Light blue ribbons represent 2.5th and 97.5th quantiles of distribution of individual variation, with shades representing median, 25th-75th quantiles and 5th-90th quantiles of posterior sample. Black points represent development rate estimates from actual observations.

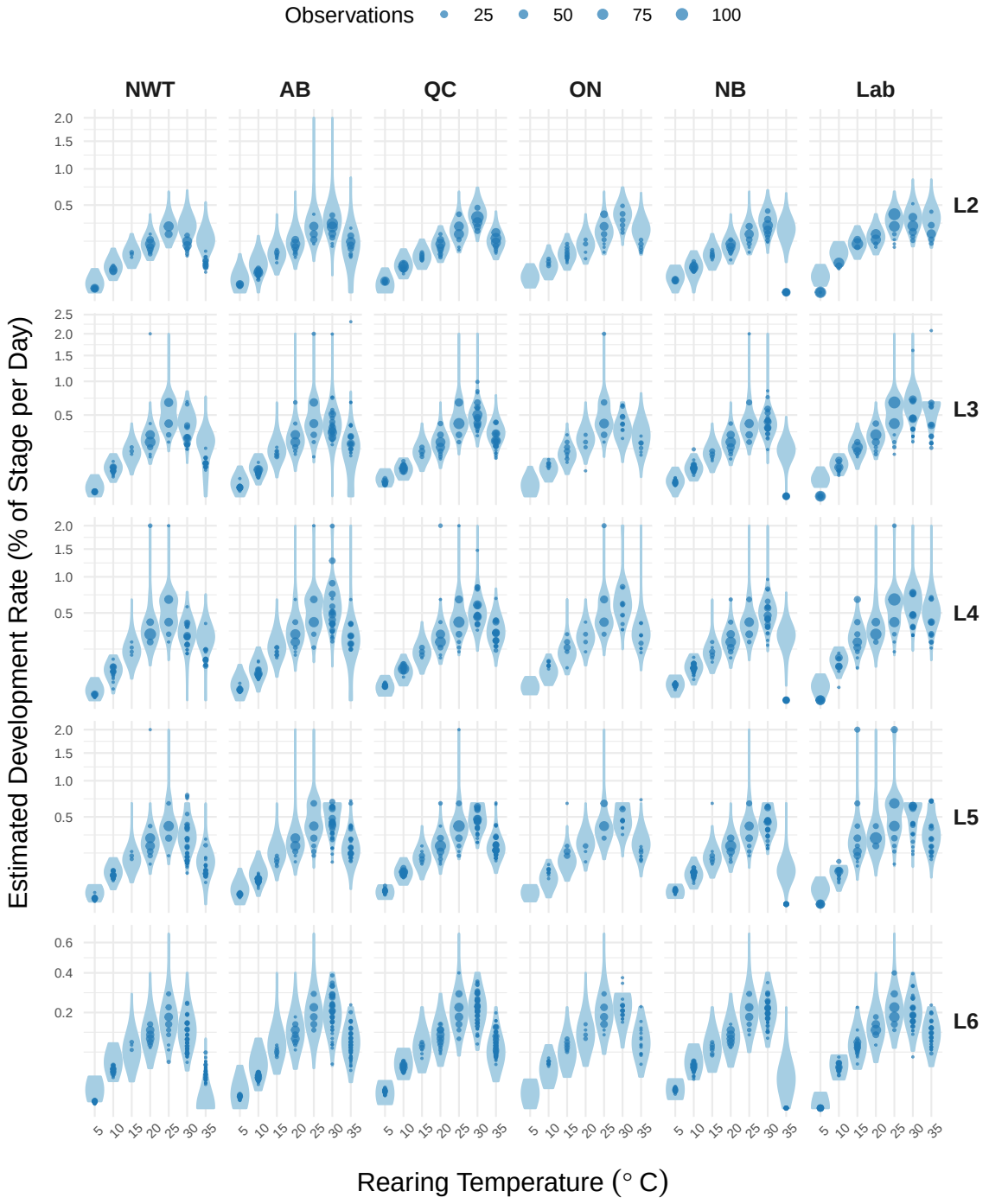


Figure 3.3: Comparison of simulations from posterior generative model to observations. Violins represent densities of simulated observations, while points represent actual observations.

temperatures, they could be partially due to the estimation procedure for development rates from the observed data.

To compensate for any discrepancies resulting from this type of estimation, sample datasets were randomly generated from the posterior distribution and interval censored as though they were observed in the laboratory following the experimental protocol, and development rates were estimated from these values as they were for the observed data. At each level, 1,000 samples were taken from the posterior distributions and 100 data points were randomly generated from each sample at each temperature. Violin plots represent the distribution of values in a group by plotting a symmetric version of the kernel density estimate for the points within a group. The violins in Figure 3.3 represent the densities of these simulated datasets, and the dark blue points show the estimated rates from the actual observations. The densities tend to capture the points well; the lack of fit in the L2 stage for Québec, Northwest Territories and Ontario is less pronounced and for many levels the extreme-temperature observations are captured just as well as those from more optimal temperatures.

A major discrepancy that is still visible is that of the New Brunswick 35°C observations and their estimates. For each development stage, the estimated development at 35°C for the New Brunswick colony is zero, but the model does not reflect that until the L6 stage. The other visible discrepancy is in the 5°C observations in the lab-reared colony. Similar to the New Brunswick 35° observations, the observed estimates are zero but the model tends higher.

Most observations fall within reasonable ranges of the violins, but in certain areas the observed values lie consistently in the tails. This occurrence is most visible in the 30 and 35°C violins of the Northwest Territories colony, and the 35°C violins of the New Brunswick and Alberta colonies, which indicates that there may still be some unexplained lack of fit in the model.

3.2.1 Influence of σ_υ Prior

Values of υ that are far from one indicate a large discrepancy between the estimated development rate curve and the observed mean development rate. A large value of the variance parameter for υ, σ_υ, could therefore lead to worse fits of the development rate curve parameters to the data, since values of υ would be free to vary widely. That is, if the lack-of-fit is allowed to be large, there is less of a penalty for parameters whose corresponding development rate curve produces estimates that are far from the data. To assess to assess the effect of the σ_υ parameter, the model was fit to the L2

stage of the New Brunswick colony at four different fixed values.

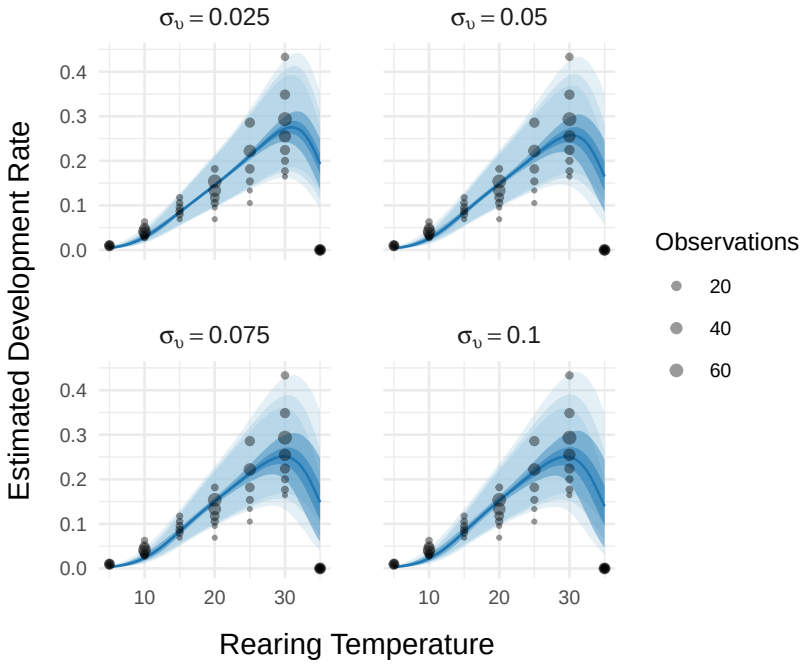


Figure 3.4: Comparison of posterior predictive plots for four fixed values of σ_υ

Figure 3.4 shows the posterior predictive plots associated with each of these fixed values, untransformed by the estimates of the lack-of-fit parameter υ. While Figure 3.2 displays splines connecting transformed estimates at each treatment temperature, Figure 3.4 displays the estimated development rate curves across the span of temperatures. If the initial hypothesis were correct, the smaller values of σ_υ would bring the bulk of posterior curves closer to the observations at each temperature. However, Figure 3.4 shows that this is not the case; in fact, the fit with large σ_υ comes closest to the point at 35°C. This indicates that the discrepancy in the fit for this population is actually likely due to either overly restrictive priors on the development rate parameters, or to an unsuitable development rate curve. Larger values of σ_υ bring central estimates closer to the observations, but also increase the variance of the development rate estimates across the posterior sample.

3.3 Comparison Across Stages

Since the life stages of individuals from a given colony are likely correlated, sharing information across stages would be a good feature of a fully realized development model. Not only does this make use

of more data in the model fitting, it could also reduce the total effective number of parameters used to fit all of the data for a given colony.

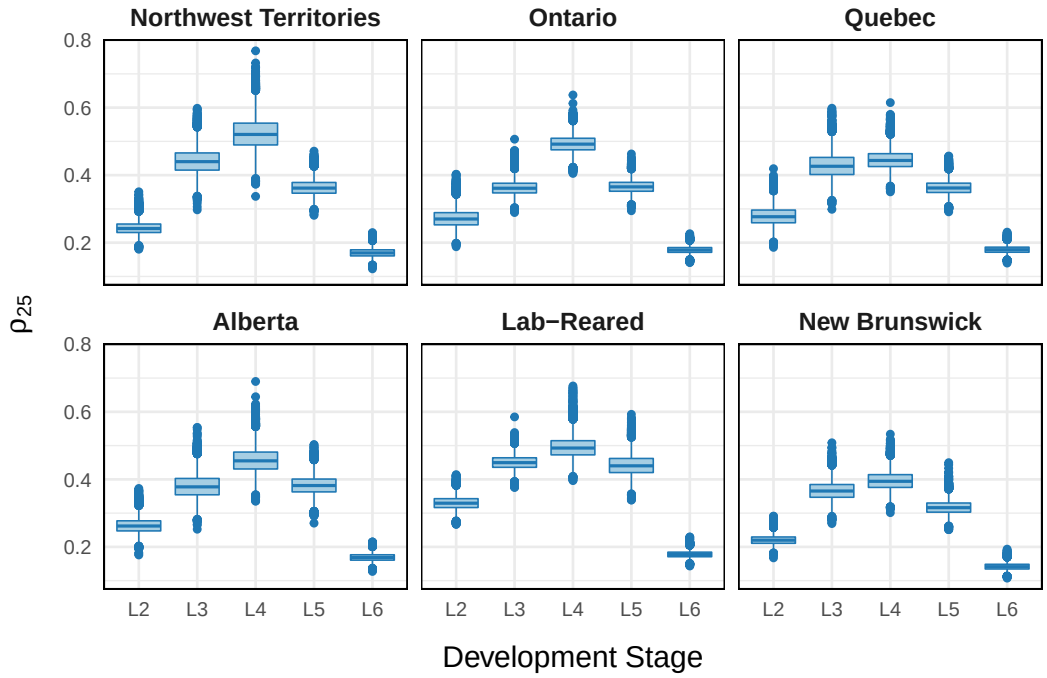


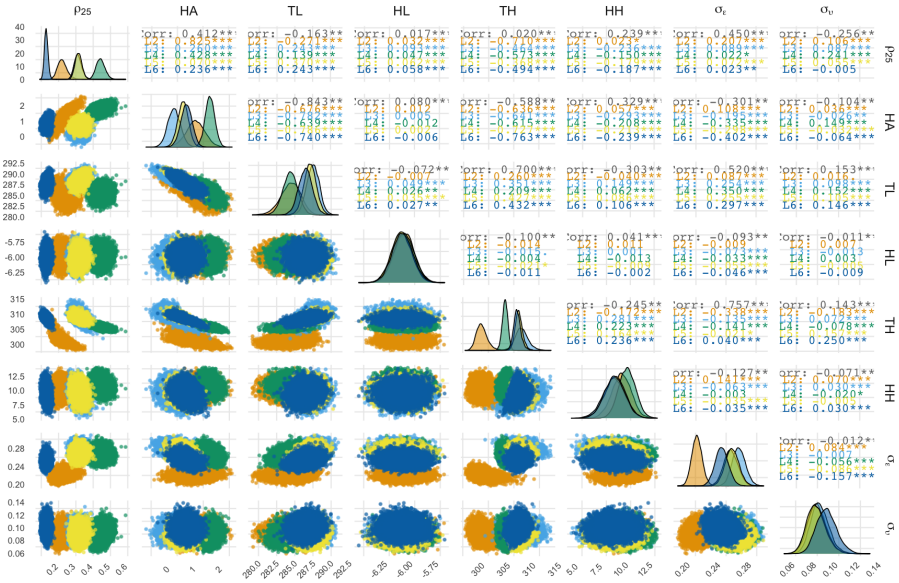
Figure 3.5: Comparison of marginal posterior distributions of ρ_{25} across stages for each colony

Box plots for each variable were generated to show any functional relationship across stages. While most parameters showed no distinguishable structure across development stages, the ρ_{25} parameter showed a strong quadratic relationship with stage (Figure 3.5). This makes sense intuitively, since ρ_{25} is effectively an amplitude parameter for development rate; it represents the pattern in the durations of each stage.

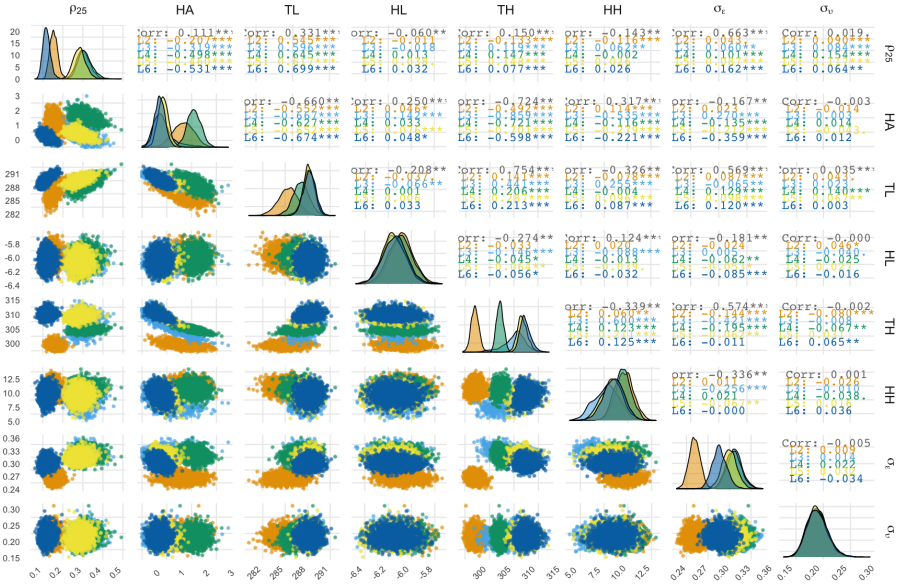
3.3.1 Structured Model

The structured model described in Chapter 2 was fitted for each colony separately. The posterior samples for each colony show no clear evidence of non-convergence, and the effective samples sizes are sufficiently large. Divergent transitions appeared in some of these model runs, but increasing the prior of σ_ϵ as in the independent model runs eliminated them.

Figure 3.6 compares the posterior distributions of the independent and structured model runs for the Ontario colony. In the pairs plot of the structured model, the ρ_{25} parameter samples have been



(a) Independent Runs



(b) Structured Model

Figure 3.6: Posterior samples for the Ontario colony from (a) independent model and (b) stage-structured model runs

estimated from the samples of x_v, y_v and ϕ. The marginal densities across stages are fairly consistent between the two model runs, but it is evident that the correlation structure of the ρ_{25} parameter is different. The correlation between ρ_{25} and H_A switches signs for each stage and decreases in

magnitude, while the correlation between ρ_{25} and T_L becomes more pronounced and the correlation between ρ_{25} and T_H effectively vanishes.

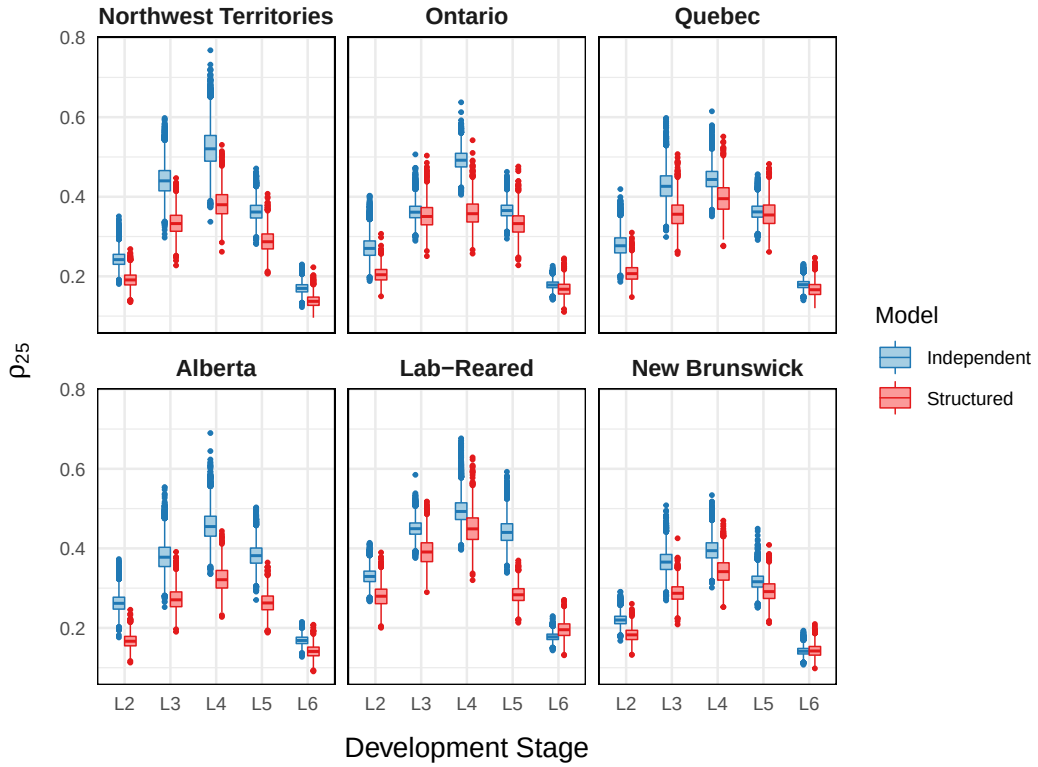


Figure 3.7: Comparison of marginal posterior distributions of ρ_{25} from unstructured model and each stage from the structured model

Figure 3.7 shows the fitted values of ρ_{25} from the structured model compared to the fitted values of ρ_{25} for the model where each level is run independently. The estimates of this parameter are systematically smaller than those from the independent runs. This is also reflected in Figure 3.8, a comparison of the mean estimates across all posterior samples of the structured model and the independent runs for each level. The cause of this discrepancy in the structured results is unclear; there is no evidence of non-convergence in the posterior samples, and the multiplicative random effect α should allow the estimates of ρ_{25} to be effectively equal for the two model types. Figure 3.8 shows the large differences in development rate estimates between the two model types in most levels of the model, especially at high temperatures in the L2 stage. The two model types appear the most similar in the L6 stage.

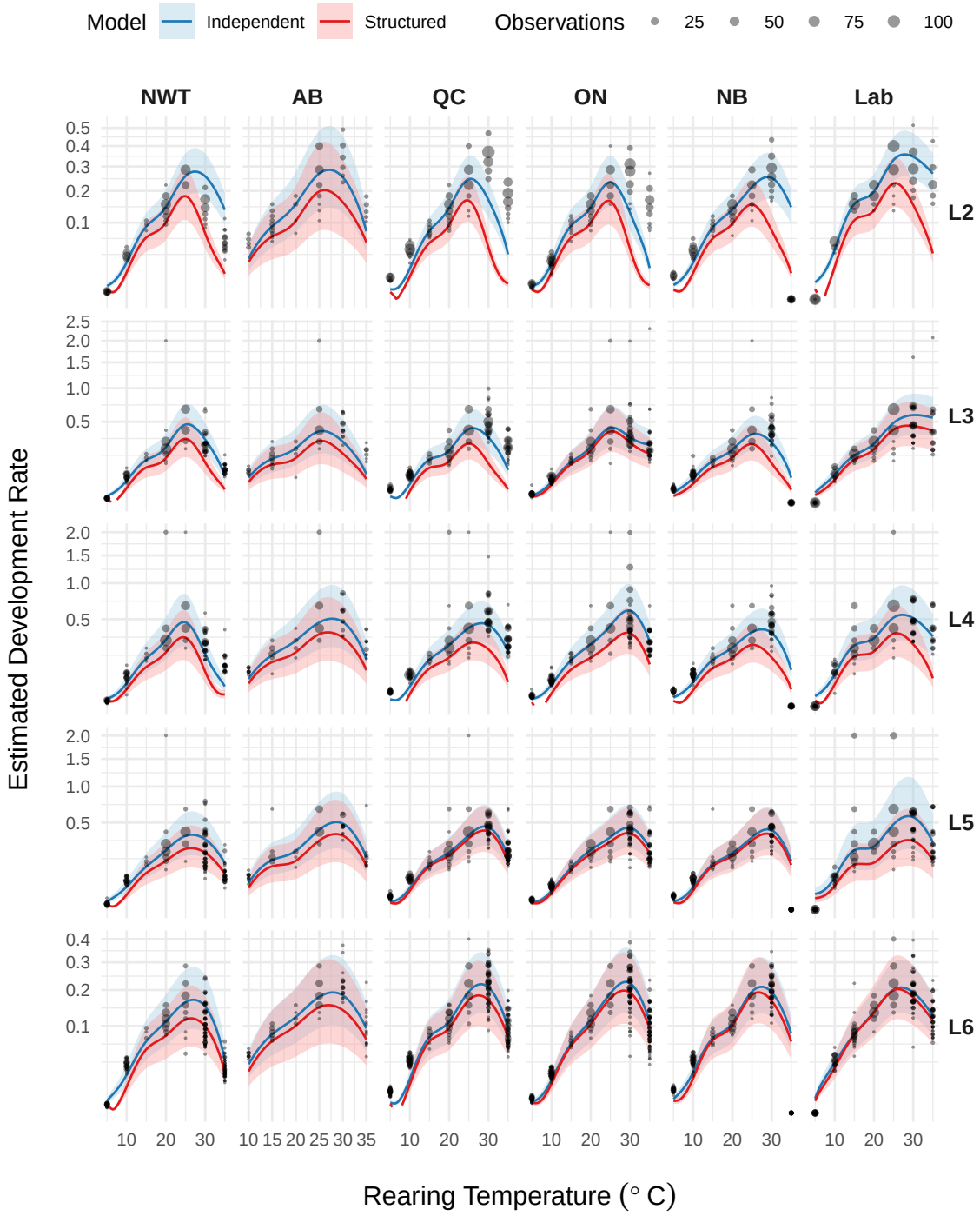


Figure 3.8: Comparison of development rate estimates at all observed temperatures between structured and unstructured models. Solid lines represent mean estimates at each observed temperature fitted with splines, while the ribbons represent the 2.5th and 97.5th percentiles of the distribution of individual variation around the mean estimates. Black points represent development rate estimates from actual observations.

3.4 Comparison Across Colonies

The behaviour of the models fitted to different colonies can be studied by comparing both the posterior distributions of the parameters and the distributions of the estimated development rates at different temperatures of interest.

3.4.1 Parameter Estimates

Figure 3.9 shows the marginal posterior samples from the unstructured model for each parameter. The quadratic shape of the ρ_{25} parameter is immediately visible across stages. The value of ρ_{25} for the lab-reared colony is consistently larger, or at the upper end, of the values of this parameter across the rest of the colonies. As previously mentioned, this parameter represents the amplitude of the development rate curve; this indicates that the insects from this colony tend to develop more quickly than the rest. This phenomenon could be explained by the colony's genetic history of being reared under similar conditions. The lab-reared colony is also consistently on the upper end of the T_H parameter estimates. This parameter represents the temperature at which half of the control enzymes have been deactivated due to high temperature denaturation, so large values of this parameter indicate tolerance to higher temperatures. The colonies in the box plots are sorted according from highest to lowest latitude of geographic origin; this gradient is most pronounced in the T_H parameter, specifically in the L2, L5 and L6 stages. The order of tolerance to high temperatures is flipped from L2 to L5; Ontario has minimal high-temperature tolerance in L2 and the more northern colonies increase by latitude, while the reverse is true in the L5 stage. In the L6 stage, tolerance to high temperatures decreases with increasing latitude. This behaviour makes sense, since we would expect colonies from higher latitudes to have less high-temperature exposure, especially during the L6 stage which typically occurs mid summer. The values of H_L and H_H, which are proportional to the slopes of the development rate curve at T_L and T_H respectively, are relatively consistent across colonies. On the other hand the H_A parameter, which has the same interpretation for 25°C, varies widely across colonies and within colonies across stages. The Alberta colony has more individual variation in development rates than other colonies, as reflected in the σ_ϵ parameter. This colony also has consistently low values of T_L, indicating a tolerance to low temperatures. The parameter estimates are, unexpectedly, consistently lower than those for the Northwest Territories population. Another anomaly in the T_L parameter is the sudden downward spike in the estimates

for the lab-reared population in the L5 stage. In the lack-of-fit parameter σ_v, the estimates are relatively consistent across province and stage until stage L6, where they increase sharply for the Northwest Territories, Québec and New Brunswick colonies.

3.4.2 Development Rate Estimates

Figure 3.10 shows violins representing the distributions of central development time estimates for each posterior draw across colonies and stages; estimates were obtained at each temperature for each posterior draw. The estimates for the lab-reared population are consistently large relative to the estimates for the other colonies, as was anticipated by the estimates of ρ_{25} for that colony. This discrepancy is especially visible in the L3 development stage, in which the decrease in development rates due to high temperatures is significantly slower for this colony than for the rest.

Across all colonies, the development rate estimates at the 30 and 35°C rearing temperatures have much higher variability than the rest of the parameters, especially in the first three stages of development; in the L2 stage, the distributions of estimates for certain colonies span the height of the development curve. The difference between the Northwest Territories colony and the others is most obvious in the high temperature estimates for stages L4-L6; the estimated development rates decrease more rapidly in the Northwest Territories colony, indicating that the insects from that colony are more negatively affected by high temperatures in late stages.

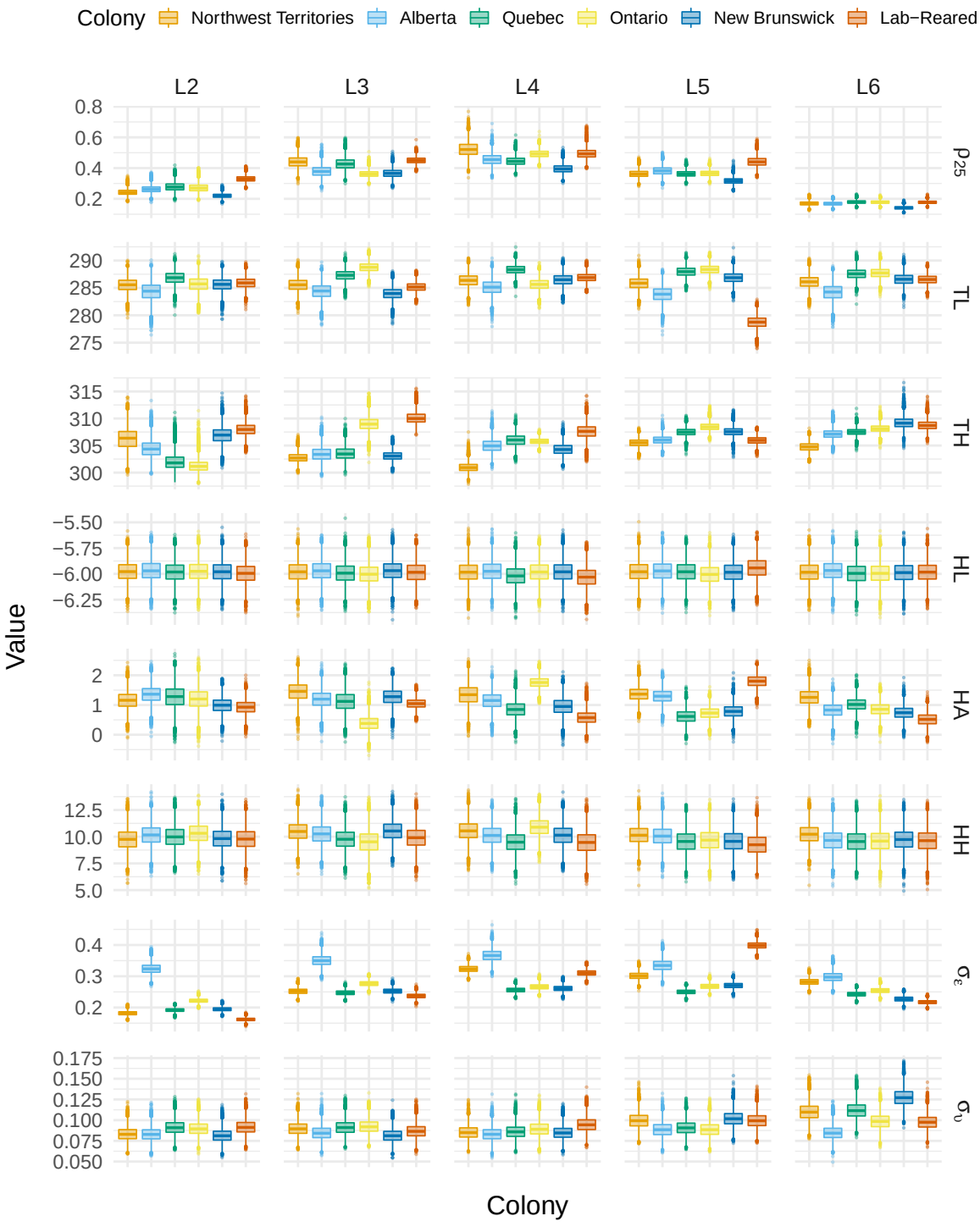


Figure 3.9: Comparison of marginal posterior distributions of each parameter from unstructured model across colonies and stages

Figure 3.10: Comparison of distribution of central development rate estimates at each rearing temperature across colonies and stages.

Chapter 4

Discussion

4.1 Convergence

The Bayesian hierarchical model was fitted successfully to every level of the dataset, with no obvious lack of convergence. Divergent transitions in the initial round of posterior sampling revealed regions of high curvature in the target distribution, despite the de-centralization of the hierarchical parameters of the model. The cause of this curvature was attributed to interval censored data; this was dealt with by shifting the prior distribution of the variance parameter for the distribution of individual variation to concentrate around larger values.

Altering the prior on that variable superficially fixed the issue at hand, but required an iterative, trial and error approach to do so. A more effective and more principled way to handle this issue of curvature would be to parameterize the model in a way that would prevent it altogether. A possibility would be to implement a prior with thicker tails on the problematic variance parameter, such as a Student-t prior, thereby increasing the distance between the two evaluations of the CDF for points far from the bulk of the distribution. Another solution, which would have major implications on data collection, would be to rely on an observation of growth such as weight or size rather than an observation of whether or not a life event occurred. This would eliminate the interval censoring, thus eliminating the cause of the problematic target surface. While this would add significant time and resource requirements to the data collection process, there are other reasons why such continuous measurements provide advantages over the discrete observations of times to moult (Section 4.3).

4.2 Model Fits

The model fits most levels of the data quite well. There was some difficulty for observations with estimated development rates close to zero but most points fell within the range of the simulated data points from the generative model (Figure 3.3). Some observations, especially at the high extremes of rearing temperature, landed in the tails of the simulated densities, indicating some lack of fit at those temperatures. Reasons for this lack of fit could be overly restrictive priors on some of the parameters of the development curve, limiting its flexibility, or an unsuitability of the selected development rate curve to fit the data. The generalization of the same development curve as presented in Ikemoto (2005) could be worth investigating, as the curve in its current form is restricted in the reference temperature for the amplitude parameter, ρ_{25}. While this temperature is said in Schoolfield et al. (1981) to be constant within a given species, the wide geographic variation in the spruce budworm and the differences in fits across colonies suggest that this parameter could also vary across colonies.

Fitting a Bayesian hierarchical model to the individual-level data is beneficial for prediction purposes. It provides a built-in generative model for simulating the population response to different climate regimes, propagating uncertainty throughout. The posterior samples also provide credible intervals for parameter values and for any function of the parameter values. The limitation of using these fits to predict the behaviour of insect populations in the wild is that they assume that the response of insects in the wild mimics the response of insects reared in laboratory conditions, under constant temperature regimes. This bias can be evaluated by comparing the estimates from the model fits to observations of wild populations. This type of validation would be advisable before accepting long term predictions based on the fits from these models. Future work to be done also includes rearing insects under a variety of fluctuating temperature regimes to assess their effect on insect development.

As shown in Figure 2.1, some treatment levels had very small sample sizes compared to others. Specifically, the 15°C treatment level showed high mortality across most colonies, and the Alberta colony contained far fewer samples than any other colony. In Bayesian models, smaller sample sizes result in larger influence of prior distributions. For this reason, the estimates from the Alberta colony may tend closer to the prior distributions than the estimates for the rest of the colonies. Future work includes modelling individuals' survival across development stages, and fitting a model using the right censored datasets containing every individual.

4.3 Comparison Across Stages

The independent model runs revealed a strong quadratic structure in the amplitude parameter of the development rate curve across stages in each colony. This led to the implementation of a model incorporating this quadratic structure across stages; the other parameters were allowed to vary independently across stages, but the ρ_{25} parameter was represented by a quadratic function of the discrete developmental stage. A random effect multiplier was included to allow for stage-specific deviation from the quadratic curve. Fitting this model resulted in different estimates than the independent runs; the estimates of ρ_{25} were systematically smaller in the structured model, and the correlation structure of the posterior was changed.

A limitation of the implementation of the structured model is that stages are treated as discrete, evenly spaced, ordinal values. This representation is an oversimplification of the way development stages actually occur. An individual is determined to have completed a stage upon moulting, and the moults themselves are marked when individuals shed their head capsules. It is possible that these occurrences are more closely related to the individuals' sizes than to an abrupt shift in their physiology causing a change in their response to temperature. It is likely that there is a gradual shift in the response over time, as an individual grows. The quadratic structure observed in the amplitude of development rate curves indicates that there is some continuum in the thermal response of individuals as they grow, so continuous observations of development such as size or weight over time could be mapped to this structure to produce a more sophisticated model of thermal response across development.

Another component that could be added to stage-structured model is a notion of correlation across stages for individuals. The current formulation of the model operates under the assumption that individuals' development relative to the population mean is independent across stages, but this assumption has never been validated. Estimates of correlation parameters from a model with this structure would contain such information. This structure would take full advantage of the individual-based data, and would provide information that could not be ascertained from cohort-based data collection.

4.4 Comparison Across Colonies

Comparing the estimates from the unstructured model across colonies revealed that the lab-reared colony produced development rate estimates that consistently stood out from the rest. The estimates indicate that insects from this colony tend to develop more quickly than those from wild populations. This suggests that estimating development rates from colonies that are not adapted for outdoor conditions such as fluctuating or extreme temperatures and natural diet can result in bias.

Comparing estimates also revealed differences in the fits of development rate curves across colonies. The full implications of these differences cannot be assessed only by comparing parameter estimates; rather, a more thorough investigation such as predictive simulations should be done to further examine the differences in thermal response. An example of such a predictive simulation would be to generate populations from the posterior distributions for every colony and simulate the populations' development over historical temperature observations. The resulting distributions of the proportions of populations in each stage over time could then be compared across colonies to assess how different colonies would fare under similar weather conditions. Future work also includes mapping development rate parameters to spruce budworm subpopulations, as determined by their genomic structure.

Bibliography

Ashcroft, M., Casanova-Katny, A., Mengersen, K., Rosenstiel, T., Turnbull, J., Wasley, J., Waterman, M., Zúñiga, G., and Robinson, S. (2016). Bayesian methods for comparing species physiological and ecological response curves. *Ecological Informatics*, 34:35–43.

Bellows Jr., T. and Birley, M. (1981). Estimating developmental and mortality rates and stage recruitment from insect stage-frequency data. *Population Ecology*, 23(2):232–244.

Brière, J.-F., Pracros, P., Roux, A.-Y. L., and Pierre, J.-S. (1999). A novel rate model of temperature-dependent development for arthropods. *Environmental Entomology*, 28(1):22–29.

Candau, J.-N., Dedes, J., Lovelace, A., MacQuarrie, C., Perrault, K., Roe, A., Studens, K., and Wardlaw, A. (2019). Validation of a spruce budworm phenology model across environmental and genetic gradients: applications for budworm control and climate change predictions.

Chuine, I. and Régnière, J. (2017). Process-based models of phenology for plants and animals. *Annual Review of Ecology, Evolution, and Systematics*, 48:159–182.

Corkrey, R., Olley, J., Ratkowsky, D., McMeekin, T., and Ross, T. (2012). Universality of thermodynamic constants governing biological growth rates. *PLoS One*, 7(2):e32003.

Damos, P. and Savopoulou-Soultani, M. (2012). Temperature-driven models for insect development and vital thermal requirements. *Psyche*, 2012. Article ID 123405.

Davidson, J. (1944). On the relationship between temperature and rate of development of insects at constant temperatures. *The Journal of Animal Ecology*, pages 26–38.

DeAngelis, D. (2018). *Individual-based models and approaches in ecology: populations, communities and ecosystems*. CRC Press.

Dempster, J. (1961). The analysis of data obtained by regular sampling of an insect population. *The Journal of Animal Ecology*, 30(2):429–432.

Dennis, B., Kemp, W., and Beckwith, R. (1986). Stochastic model of insect phenology: estimation and testing. *Environmental Entomology*, 15(3):540–546.

Feng, X., Liang, Y., Gallardo, B., and Papeş, M. (2020). Physiology in ecological niche modeling: using zebra mussel's upper thermal tolerance to refine model predictions through Bayesian analysis. *Ecography*, 43(2):270–282.

Fu, Y., Campioli, M., Oijen, M. V., Deckmyn, G., and Janssens, I. (2012). Bayesian comparison of six different temperature-based budburst models for four temperate tree species. *Ecological Modelling*, 230:92–100.

Gelman, A., Rubin, D., et al. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, 7(4):457–472.

Hilbert, D. and Logan, J. (1983). Empirical model of nymphal development for the migratory grasshopper, *Melanoplus sanguinipes* (Orthoptera: Acrididae). *Environmental Entomology*, 12(1):1–5.

Hoffman, M. and Gelman, A. (2014). The No-U-Turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo. *Journal of Machine Learning Research*, 15(1):1593–1623.

Holling, C. (1965). The functional response of predators to prey density and its role in mimicry and population regulation. *The Memoirs of the Entomological Society of Canada*, 97(S45):5–60.

Hudes, E. (January 1982). *A method for statistically estimating insect phenology and its application to the spruce budworm*. Cornell University. [Unpublished master's thesis].

Hudes, E. and Shoemaker, C. (1988). Inferential method for modeling insect phenology and its application to the spruce budworm (Lepidoptera: Tortricidae). *Environmental Entomology*, 17(1):97–108.

Hultin, E. (1955). The influence of temperature on the rate of enzymic processes. *Acta Chemica Scandinavica*, 9(170):r1710.

Iizumi, T., Yokozawa, M., and Nishimori, M. (2009). Parameter estimation and uncertainty analysis of a large-scale crop model for paddy rice: Application of a Bayesian approach. *Agricultural and Forest Meteorology*, 149(2):333–348.

Ikemoto, T. (2005). Intrinsic optimum temperature for development of insects and mites. *Environmental Entomology*, 34(6):1377–1387.

Ikemoto, T. and Takai, K. (2000). A new linearized formula for the law of total effective temperature and the evaluation of line-fitting methods with both variables subject to error. *Environmental Entomology*, 29(4):671–682.

Janisch, E. (1932). The influence of temperature on the life-history of insects. *Transactions of the Royal Entomological Society of London*, 80:137–168.

Johnson, F. and Lewin, I. (1946). The growth rate of *E. coli* in relation to temperature, quinine and coenzyme. *Journal of Cellular and Comparative Physiology*, 28(1):47–75.

Kempton, R. (1979). Statistical analysis of frequency data obtained from sampling an insect population grouped by stages. In Ord, J., Patil, G., and Taillie, C., editors, *Statistical distributions in ecological work*, pages 401–418. International Cooperative Publishing House.

Lactin, D., Holliday, N., Johnson, D., and Craigen, R. (1995). Improved rate model of temperature-dependent development by arthropods. *Environmental Entomology*, 24(1):68–75.

Logan, J., Wollkind, D., Hoyt, S., and Tanigoshi, L. (1976). An analytic model for description of temperature dependent rate phenomena in arthropods. *Environmental Entomology*, 5(6):1133–1140.

Lysyk, T. (1989). Stochastic model of eastern spruce budworm (Lepidoptera: Tortricidae) phenology on white spruce and balsam fir. *Journal of Economic Entomology*, 82(4):1161–1168.

Marquardt, D. (1963). An algorithm for least-squares estimation of nonlinear parameters. *Journal of the Society for Industrial and Applied Mathematics*, 11(2):431–441.

Monnahan, C. and Kristensen, K. (2018). No-U-turn sampling for fast Bayesian inference in ADMB and TMB: Introducing the adnuts and tmbstan R packages. *PloS one*, 13(5).

Neal, R. (2003). Slice sampling. *Annals of Statistics*, 31(3):705–741.

Osawa, A., Shoemaker, C., and Stedinger, J. (1983). A stochastic model of balsam fir bud phenology utilizing maximum likelihood parameter estimation. *Forest Science*, 29(3):478–490.

Perrault, K., Wardlaw, A., Candau, J.-N., Irwin, C., Demidovich, M., MacQuarrie, C., and Roe, A. (2021). From branch to bench: establishing wild spruce budworm populations into laboratory colonies for the exploration of local adaptation and plasticity. *The Canadian Entomologist*, page 1–17. Advanced online publication.

R Core Team (2020). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Read, K. and Ashford, J. (1968). A system of models for the life cycle of a biological organism. *Biometrika*, 55(1):211–221.

Rebaudo, F. and Rabhi, V.-B. (2018). Modeling temperature-dependent development rate and phenology in insects: review of major developments, challenges, and future directions. *Entomologia Experimentalis et Applicata*, 166(8):607–617.

Régnière, J. (1987). Temperature-dependent development of eggs and larvae of *Choristoneura fumiferana* (Clem.) (Lepidoptera: Tortricidae) and simulation of its seasonal history. *The Canadian Entomologist*, 119(7-8):717–728.

Régnière, J. and Powell, J. (2013). Animal life cycle models (Poikilotherms). In Schwartz, M., editor, *Phenology: An Integrative Environmental Science*, pages 295–316. Springer.

Régnière, J., Powell, J., Bentz, B., and Nealis, V. (2012). Effects of temperature on development, survival and reproduction of insects: experimental design, data analysis and modeling. *Journal of Insect Physiology*, 58(5):634–647.

Régnière, J., Saint-Amant, R., Béchard, A., and Moutaoufik, A. (2014). *BioSIM 10: User's manual*. Laurentian Forestry Centre.

Schoolfield, R., Sharpe, P., and Magnuson, C. (1981). Non-linear regression of biological temperature-dependent rate models based on absolute reaction-rate theory. *Journal of Theoretical Biology*, 88(4):719–731.

Sharpe, P., Curry, G., DeMichele, D., and Cole, C. (1977). Distribution model of organism development times. *Journal of Theoretical Biology*, 66(1):21–38.

Sharpe, P. and DeMichele, D. (1977). Reaction kinetics of poikilotherm development. *Journal of Theoretical Biology*, 64(4):649–670.

Southwood, T. and Jepson, W. (1962). Studies on the populations of *Oscinella frit* L.(Dipt: Chloropidae) in the oat crop. *The Journal of Animal Ecology*, 31(3):481–495.

Stan Development Team (2020). *Stan Modeling Language Users Guide and Reference Manual 2.25*.

Stedinger, J., Shoemaker, C., and Tenga, R. (1985). A stochastic model of insect phenology for a population with spatially variable development rates. *Biometrics*, 41(3):691–701.

Stinner, R., Gutierrez, A., and Butler, G. (1974). An algorithm for temperature-dependent growth rate simulation. *The Canadian Entomologist*, 106(5):519–524.

Tanigoshi, L. and Logan, J. (1979). Tetranychid development under variable temperature regimes. *Recent Advances in Acarology*, 1:165–175.

Taylor, F. (1981). Ecology and evolution of physiological time in insects. *The American Naturalist*, 117(1):1–23.

Thorsen, S. and Höglind, M. (2010). Modelling cold hardening and dehardening in timothy. Sensitivity analysis and Bayesian model comparison. *Agricultural and Forest Meteorology*, 150(12):1529–1542.

Uvarov, B. et al. (1931). Insects and climate. *Transactions of the Royal Entomological Society of London*, 79(pt. 1).

Vehtari, A., Gelman, A., Simpson, D., Carpenter, B., and Bürkner, P.-C. (2021). Rank-Normalization, Folding, and Localization: An Improved $\widehat{R}$ for Assessing Convergence of MCMC. *Bayesian Analysis*, -1(-1):1 – 28.

Weber, J., Volney, W., and Spence, J. (1999). Intrinsic development rate of spruce budworm (Lepidoptera: Tortricidae) across a gradient of latitude. *Environmental Entomology*, 28(2):224–232.

Worner, S. (1992). Performance of phenological models under variable temperature regimes: consequences of the Kaufmann or rate summation effect. *Environmental Entomology*, 21(4):689–699.