

A Mechanistic Model of *Drosophila suzukii* Population Dynamics In  
Contemporary and Future Climates

by

Brent Aaron Bradley Langille

A Thesis  
presented to  
The University of Guelph

In partial fulfilment of requirements  
for the degree of  
Doctor of Philosophy  
in  
Environmental Science

Guelph, Ontario, Canada

© Aaron Langille, April, 2017

# Abstract

Brent AB. Langille  
University of Guelph, 2017

Advisor:  
Jonathan A. Newman

*Drosophila suzukii* is a relatively new threat to the soft-skinned fruit industry in North America. The presence of this pest in North America is a concern and assessing the risk of potential infestation and damage can help guide regional management strategies. In this work, I present a mechanistically-realistic population model parameterized with empirical data based on laboratory observations. The principle environmental driver of vital rates for the model is temperature though I include considerations for reproductive diapause and quality of fruit available. Relative population estimates are produced for important berry producing locations using contemporary temperature data. Southern regions appear to be most suitable for supporting high population densities due to warm temperatures throughout the year though some temperate regions also may be at risk.

Understanding *D. suzukii*'s potential to shift in abundance and range due to changing climate is an important part of long term mitigation and management strategy. The population dynamics model is run using temperature data derived from several Global Circulation Models (CMIP5) with a range of relative concentration pathway (RCP) predictions. Mean consensus between the models suggest that without adaptation to both higher prolonged temperatures and higher short-term temperature events *D. suzukii* population levels are likely to drop in currently higher-risk regions. Northern latitudes may experience increased populations due to milder winter and more developmentally-ideal summer conditions though many of these regions are not currently known for soft-skinned fruit production.

Finally, effective communication of model design, implementation and execution details are of genuine importance. These details are provided using a modified version of the standardized Overview, Design Concepts and Details protocol proposed by Grimm et al. (2006, 2010) in hopes that the *D. suzukii* population dynamics model may be used or extended by third parties. A discussion of parsimony as it relates to mechanistically-rich models in the context of modelling objectives is presented as is a brief discussion of model validation. I conclude with a "user manual" for the population dynamics model and a brief overview of a web-based version.

# Acknowledgements

Thank you to Jonathan Newman for agreeing to supervise and support this work. Also, for knowing exactly when to assist and when to leave me to find the answers on my own (i.e., developmentally appropriate ignoring).

Thank you to my family who stood beside me through this colossal journey during equal measures of productivity, procrastination, mania and questioning-it-all.

Thank you to my University of Guelph lab mates Gerry Ryan and Lisa Emiljanowicz. Your efforts paved the way for this work and it wouldn't be here otherwise.

Thank you to my Laurentian University lab mate Ellen Artega. May all your research partners be just as keen and as knowledgeable as you are.

Thank you to my advisory committee members, Andy Alstadt, Kim Cuddington, and Thomas Merritt. Your feedback made this work better.

Thank you to my entire Laurentian University family for having my back every step of the way.

# Table of Contents

Abstract .....	ii
Acknowledgements .....	iii
Table of Contents .....	iv
List of Tables .....	vi
List of Figures .....	vii
1. Modelling <i>Drosophila suzukii</i> population dynamics - Introduction and Review .....	1
1.1 Introduction .....	1
1.2 <i>Drosophila suzukii</i> .....	1
1.3 Primer on Modelling Philosophy and Approaches .....	2
1.4 Modelling Invasive Insect Pests .....	5
1.5 Modelling <i>D. suzukii</i> .....	6
2. North American Invasion of Spotted-Wing <i>Drosophila</i> ( <i>Drosophila suzukii</i> ): A Mechanistic Model of Population Dynamics .....	9
2.1 Abstract .....	9
2.2 Introduction .....	10
2.2.1 <i>D. suzukii</i> Life History .....	11
2.3 Model .....	13
2.3.1 Fecundity Rate .....	14
2.3.2 Juvenile Development Rate .....	15
2.3.3 Temperature-Dependent Mortality .....	16
2.3.4 Effect of Fruit Quality on Development and Mortality .....	17
2.3.5 Model State Equations .....	20
2.3.6 Environmental Equations .....	22
2.3.7 Model Implementation .....	24
2.4 Results .....	25
2.4.1 Base Model Dynamics .....	25
2.4.2 Effects of Starting Population Size and Appearance Timing .....	26
2.4.3 Diapause Sub Model .....	27
2.4.4 Observed Temperature Data simulations .....	30
2.4.5 Fruit Sub Model .....	31
2.4.6 The Complete Model .....	33
2.4.7 Extrinsic Mortality .....	35
2.5 Discussion .....	36
2.6 Acknowledgements .....	40
3. The Impacts of Climate Change on the Abundance and Distribution of the Spotted Wing <i>Drosophila</i> ( <i>Drosophila suzukii</i> ) in The United States and Canada .....	41
3.1 Abstract .....	41
3.2 Introduction .....	42

3.3	Materials and Methods .....	44
3.4	Results .....	45
3.5	Discussion.....	52
3.6	Acknowledgements.....	55
4.	Model Design and Use .....	56
4.1	Abstract .....	56
4.2	Introduction.....	56
4.3	Overview, Design Concepts and Details (ODD) .....	57
4.3.1	Purpose.....	57
4.3.2	State Variables, Scale and Entities.....	57
4.3.3	Process Overview and Scheduling .....	59
4.3.4	Design Concepts .....	60
4.3.5	Initialization, Input Data and Sub models .....	69
4.4	Model Technical Details .....	70
4.4.1	Compiling the Models .....	70
4.4.2	Executing the Models .....	72
4.4.3	Multiple Simulation Model.....	75
4.4.4	Web-based Model .....	77
4.5	Discussion .....	79
5	General Discussion and Conclusion .....	80
5.1	Introduction.....	80
5.2	Mechanistically Modelling D. Suzukii Population Dynamics .....	80
5.2.1	Identifying D. Suzukii Knowledge Gaps .....	81
5.2.2	Model validation.....	83
5.3	Exploring the Implications of Climatic Change .....	84
5.4	Improvement Through Collaboration and Open Access .....	86
6	Bibliography.....	89
	Appendix 1: Chapter 2 Supplemental Figures and Model Comparison .....	98
	Appendix 2: Table of Model Parameters.....	104
	Appendix 3: Individual GCM-RCP Population Maps .....	106

## List of Tables

Table 2.1 Estimated parameters that vary by stage for temperature-dependent mortality (see e.g. Emiljanowicz et al. 2014). The $\beta_{i,0}$ parameters are dimensionless shape parameters, and the $\mu_{i,\max}$ are per capita daily mortality rates.....	17
Table 2.2 Parameters used to generate model temperature profiles per eqn 23.....	23
Table 2.3 Latitude and longitude of 12 of important blueberry and/or strawberry producing counties and townships in the United States and Canada. Sorted by latitude (north to south). See Figure A1.3 in the supplemental material for a map of these locations.....	24
Table 3.1 Official contributor name, institute ID and model name for the GCMs used herein.....	44
Table 3.2 Key base model parameters for simulations, cf. Langille et al. (2016, supplemental) for complete parameter value listing. .....	45
Table 4.1 Summary of files and directories present in the initial source code release. ....	71
Table 4.2 Parameter names, values and descriptions required for the singleSim executable. ....	73
Table 4.3 Code that can be altered to customize the multiSim or multiSimDemo execution. Any code modifications require a re-compile before execution.....	76
Table A1.1 Model result maxima used for normalization and sub model inclusion state per figure. All simulations were started with 100 females unless otherwise stated.....	103
Table A1.2 Diapause sub model parameters constant for all simulations unless otherwise stated. ....	103
Table A1.3 Fruit sub model parameters constant for all simulations unless otherwise stated...103	103
Table A2.1 Model parameters including values, units of measure and supplemental notes on estimation methods or purpose.....	103

## List of Figures

Figure 2.1 Schematic of the modelled population stage structure and relationships between state variable equations including principal input and output parameters. Most development and mortality processes are temperature (T) dependent and rely on the fruit sub model (fr) when included. Here mortality is represented as the sum of intrinsic and extrinsic factors (see main text) .....	13
Figure 2.2 a) Fruit quality development time as described in eqn 8 for a simulated temperature profile (see section 2.6.2) with a harvest delay, $\omega$ , 50. b) illustrates fruit quality index (eqn 11) for influencing development and fruit quality decline (eqn 13) for influencing mortality.....	19
Figure 2.3 Female age-specific egg viability. Changes in viability correspond to $v_i$ in eqn 15 where $v_1 = 0.832$ , $v_2 = 0.807$ , $v_3 = 0.763$ , $v_4 = 0.556$ , $v_5 = 0.324$ , $v_6 = 0.257$ (see Emiljanowicz et al. 2014 for further information).....	20
Figure 2.4 Modelled temperature profiles approximating 20-year mean daily temperature for Chicoutimi QC (A), Clark county WA (B), Santa Barbara county CA (C) and Hillsborough county FL, (D).....	23
Figure 2.5 Normalized cumulative female population for various constant temperatures. Simulations are 365 time steps in length, begin with 10 fecund females and do not include the diapause or fruit quality sub models.....	25
Figure 2.6 Cumulative population sizes for varied initial introduction dates and populations sizes. In both cases, introduced flies are fecund females and the simulated timeframe uses model temperature profile Temp B. Individual simulations were run for 365 time steps. ....	26
Figure 2.7 The effect of diapause termination temperature on cumulative populations for model temperature profiles Temp A through D. Each simulation starts with 10 fecund females introduced into the simulation at time step 0 and was run for 365 days. Diapause induction is fixed at 10 hours of daylight for all simulations. Note the difference in the y-axis scaling between graphs. ....	28
Figure 2.8 Normalized per-day female count for each of the model temperature profiles. 10 fecund females are introduced into the simulation on the time step (starred) for which diapause would be terminated due to adequate temperature. ....	29
Figure 2.9 Normalized per-day female (all adult stages) population dynamics using observed daily mean temperature data for Clark County, WA (1993-2013). Bold line indicates simulation using 20 year average daily temperature. 10 females are introduced into the simulation on the time step where temperature is adequate to break diapause.....	30

Figure 2.10 Fruit sub model effect on cumulative population. Simulations start with 10 fecund females introduced on time step 75 and are run for a duration of 365 time steps. Model temperature profile Temp B is used. Cumulative population is maximized as both $\omega$ (the quality rate limiter) and harvest lag increase resulting in optimal food availability for the simulated population.....	32
Figure 2.11 Population dynamics, including diapause and fruit quality considerations, for various soft-skin producing counties and townships across the United States and Canada. Cumulative population is normalized to show relative potential population sizes. Time step at which diapause would terminate and peak adult female population occurs for the diapause-only and diapause-fruit models are also shown. Locations are sorted by latitude (low to high).....	34
Figure 2.12 The effect of increasing extrinsic mortality in proportion to intrinsic mortality. As the extrinsic mortality increases, the cumulative population decreases.....	36
Figure 3.1 Schematic of the <i>D. suzukii</i> population dynamics model presented in Langille et al. (2016). This schematic highlights population stage structure and relationships between state variable equations including principal input and output parameters. Fecundity ( $f$ ), development ( $d$ ) and mortality ( $\mu$ ) processes are temperature ( $T$ ) dependent while development and mortality also rely on the fruit quality sub model ( $fr$ ). In addition to eggs, pupa and adult males, the model contains three juvenile instar stages and seven adult female stages in order to account for variation in egg viability ( $v$ ). $\theta$ represents the ratio of males to females and is typically set at 0.5. Solid arrows indicate developmental stage transition, dashed arrows indicate fecundity and dot-dashed arrows indicate mortality. ....	43
Figure 3.2 a) Sample temperature profiles for Burlington, New Jersey from the CCSM4 model with different RCP-date combinations. Model-RCP combinations were selected to illustrate a range of output possibilities. b) and c) illustrate fruit quality sub model response and model population based on temperature profiles in a).....	47
Figure 3.3 Aggregated consensus maps of the modelled populations. Data are the computed mean normalized total population (all life stages) from all four climate models. Normalization is based on the entire pre-mean data set. White cells located in the interior parts of the maps reflect missing values in the downscaled climate model outputs and a 0.25 degree bi-linear interpolation was applied to all data cells.....	49
Figure 3.4 (upper) Mean normalized population distribution across North America per-timeframe. Normalization occurs across the entire data set and mean is across all models and RCPs. (lower) Corresponding per-timeframe coefficient of variation across all models and RCPs. Coefficient of variation is based on the normalized population data. White cells located in the interior parts of the maps reflect missing values in the downscaled climate model outputs and a 0.25 degree bi-linear interpolation was applied to all data cells. ....	49
Figure 3.5 Grid cell coefficient of variation in population sizes versus the grid cell mean population size. The uncertainty in the population sizes that result from the differences between the climate model–RCP combinations tends to be larger as the mean of those population sizes gets smaller. Not surprisingly, the uncertainty is also larger later in the century than earlier.....	50

Figure 3.6 Normalized populations for historic temperatures and GCM projections. GCM projections are averaged over all climate model-RCP combinations where $\times$ denotes mean population, whiskers denote maximum and minimum values, interior bar marks quartiles and small circles denote outliers. For averaged historic temperatures, large circles represent population normalized to GCM results.....	51
Figure 3.7 a) and b) Normalized populations over 20 years of historic data (1993 through 2013) for Hillsborough, Florida and Fraser Valley, British Columbia respectively. The dashed line indicates the mean annual population while the bold-solid line indicates the population for the mean temperatures across all years. c) and d) illustrate the daily temperature data for all years and the mean daily temperatures across all years for Hillsborough, Florida and Fraser Valley, British Columbia respectively.....	54
Figure 4.1 Simplified UML diagram of model object relationships. All relationships are simple association except SWDCellMulti that extends SWDCellSingle .....	59
Figure 4.2 Simplified information flow diagram of the base model including diapause and fruit quality sub models. $\Theta$ determines the ratio of females to males.....	60
Figure 4.3 Two reconfigurations of the base population model presented in Chapter 2. The first reconfiguration aggregates the three larval instars into a single while the second maintains three instars but averages the entire maximum development rate across all three. For both cases fruit quality and diapaus are ignored, mortality and fecundity are set to 0 and a constant temperature of 23.5°C was used. Resulting pupal population is shown.....	65
Figure 4.4 Pupal and adult female life stage populations produced by the original and simplified version of the model for various temperature profiles. All graphs have been adjusted at the y-axis begin to highlight productive areas population change. All simulations begin with 10000 eggs at time step zero and include the diapause sub model. Fruit quality is not included.....	66
Figure 4.5 Total population (log-scale) for both versions of the model for the same temperature profiles used in the previous figure.....	67
Figure 4.6 a) Proportional difference between the original and simplified versions of the model for various temperature profiles. b) Mean annual temperature for each of the profiles. Bars indicate standard deviation. c) Mean difference in days (x) including quartiles, standard deviation (bars) and outliers (circles) between the simplified model and the original.....	68
Figure 4.7 Execution of all three compile scripts within the source directory. The final binaries are shown at the bottom of the right-most column. Example is shown using Bash shell for the Windows 10 Subsystem for Linux including the GNU 4.8.4-2ubuntu1~14.04.3 Ubuntu C++ compiler.....	72
Figure 4.8 Sample execution of singleSim using the parameters described in Table 4.2 .....	73
Figure 4.9 The first five entries for samples of both a temperature file and a parameter configuration file included with the initial release of the model.....	74

Figure 4.10 Output showing the first 10 lines and the last 20 lines of a sample simulation. ....	75
Figure 4.11 Directory structure and contents before and after execution of multiSimDemo (similar for multiSim). Directory TemperatureFiles contains properly formatted input temperature files with a naming convention that supports iterative processing. Corresponding per-time step and summary result files are created in output directory DATA. ....	77
Figure 4.12 Configuration options for each of the graph panels of the web-based simulator. Each panel can be configured independently.....	78
Figure 4.13 Four independently configured graph panels. ....	78
Figure 5.1 Daily maximum, minimum and mean temperatures across all GCM-RCP combinations for Chicoutimi QC and Wayne MS. In general, maxima and minima diverge from the mean with time. ....	85
Figure A1.1 $fT$ (eqn 1) for a range of temperature values between 0°C and 40°C. ....	98
Figure A1.2 The effect of both diapause threshold daylight hours and diapause threshold temperature on the population size. ....	98
Figure A1.3 A map of the locations in the United States and Canada for which observed temperature data was obtained and used to parameterized the presented model. ....	99
Figure A1.4 Sensitivity tests on the intrinsic mortality parameters when the base model, diapause sub model and fruit sub model are all considered. Increasing the multiplier for $\mu_{max}$ does increase the resulting population size but does not change the shape of the population growth curve over time. ....	99
Figure A1.5 (a) temperature profile for Clark WA, 2003 and (b) a simulated temperature curve (see Chapter 2 text, Figure 2.3, profile B). Fruit quality index in (a) and (b) was parameterized using gt = 4, h = 50, phi = 1 and lambda = 0.1. Stage specific population structure (c through f) for simulations using different temperature profiles (a and b). Population stages from top to bottom are females (all stages combined), males, pupae, instar 3, instar 2, instar 1 and eggs. ....	100
Figure A1.6 The model presented herein parameterized for comparison with the model of Wiman et al. (2014). Data for (a) provided by Wiman et al. for replication here. Temperature data is 2013 mean daily for Wilmington NC, Salem OR and Parlier CA. Diapause and fruit quality sub models are ignored (b). The model run for 2013 temperatures with the fruit and diapause sub models included (c). ....	102

# 1. Modelling *Drosophila suzukii* population dynamics - Introduction and Review

## 1.1 Introduction

In this introductory chapter I introduce the biological problem at the heart of this dissertation. Specifically, I am concerned with exploring the potential distribution of the invasive crop pest, *Drosophila suzukii* (commonly referred to as the spotted wing drosophila). Using a mathematical modelling approach, I consider population potential in both contemporary and future climates. To provide context for this work I discuss the use of mathematical models in ecology and then proceed to provide some background on invasive pest modelling. Finally, I introduce the idea of modelling *D. suzukii* and provide some general justification for the high-level modelling choices made in this work.

## 1.2 *Drosophila suzukii*

*Drosophila suzukii* has been a pest species in Japan since the 1930s (Kanzawa 1936) and was known to European fruit growers in 2008 (Hauser 2011). It is now believed to be present on all continents except Antarctica and reports of its growing geographic range continue to be published (Asplen et al. 2015, Lavrinenko et al. 2016, Andreazza et al. 2016). Its presence and prevalence in North America is relatively recent with sightings as early as 1980 in Hawaii and confirmation in several states and Canadian provinces by 2009 (Hauser 2011, Asplen et al. 2015).

What makes *D. suzukii* an important concern, from a biological perspective, is its preference for fresh fruit at or near optimal ripeness compared with similar pests which feed on or make use of rotting fruit. It is the female's distinct serrated ovipositor that allows it to pierce through soft skinned fruit to deposit its eggs. As the eggs hatch the juvenile stages feed on the fruit reducing its quality and bacterial infection can occur at the oviposition site reducing fruit quality even further.

Empirical estimates of *D. suzukii*-related infestation damage have yet to be published but Bolda (2010) has predicted that economic losses could be as high as \$500 million given a 20% percent crop loss due to infestation. Because of this destructive potential there has been a significant push in research on *D. suzukii* ranging from biological factors such as growth and development rates, overwintering, cold and heat tolerance, and reproductive diapause to mitigation and management strategies including identification of wild hosts, pesticide and physical barriers, and education campaigns.

### 1.3 Primer on Modelling Philosophy and Approaches

For ecological systems with extremely slow processes, extensive spatial extents or high numbers of complex interactions mathematical and computer-based models can be the only practical way to explore or estimate population dynamics (Hilborn and Mangel, 1997, Laurenroth 2003). Even when field work is feasible researchers often turn to or generate models as simplified abstractions that test their understanding of key ecological processes. By iteratively comparing modelled results and measured data, a theory-observation loop is created whereby observation helps to refine simulated processes and parameters while modelled output gauges the accuracy of the current level of understanding of such processes and guides further empirical work (Wu 1994). Baumgärtner et al. (2008) list nine principle purposes for creating models within the context of ecological economics but these principles are general enough to be applicable in most contexts:

*Theory-development*: models represent our initial understanding of the processes and functioning of a particular system. The early stages of model development and testing can help to identify areas where theory or hypotheses need to be clarified.

*Theory-testing*: completeness of general theories can be tested often by comparing (qualitatively or quantitatively) observed data with modelled results.

*Generalization*: considered together, a suite or series of models can help to identify the most general properties of a system.

*Understanding*: models can illustrate the general functioning of a system or highlight the consequences of change.

*Explanation:* similar to the previous point but with an emphasis on the mechanisms rather than the consequences of change.

*Prediction:* models may be used to predict (qualitatively or quantitatively) the future state of a system.

*Decision-support:* often an amalgamation of other purposes, models can support decision making through the illustration of alternate contemporary or future scenarios, through communication, explanation and teaching and through theory-testing.

*Communication:* Simplified and specifically-designed models can be used to present knowledge and information to stakeholders or general (not necessarily the scientific community) audiences.

*Teaching:* models, often significantly simplified (i.e., toy models), can be used as pedagogical or instructional tools. The instructional value may be in a particular system of interest, in analytical techniques or in the development and use of models.

Irrespective of the intended purpose there are several well-established approaches to model development. Phenomenological or statistical models attempt to predict future conditions, such as population size or distribution based on current or historical data. They tend to operate at a high level of abstraction as a hypothetical relationship between the variables and the data, whereby the strength of the relationship is determined by how well the data is described (Hilborn and Mangel, 1997). This abstraction occurs through assumptions and simplifications of the underlying processes which lead to the data. Examples range from simple regressions and ANOVA (Gillman 2009), to more complex Bayesian techniques (Polino et al 2007, Hobbs and Hilborn 2006, Hilborn and Mangel 1997). Conversely, mechanistic models attempt to explicitly consider the biologic and environmental processes (Hilborn and Mangel 1997, Liebold 1995). There are important advantages and disadvantages that need to be considered when choosing a modelling approach. For example, statistical models can often produce a good fit to observed data with fewer parameters (Nathan et al 2001) but do not offer much in the way of insight into the specific processes. The higher degree of abstraction favours elucidation and illustration of the general features of a system or set of systems but such general models lack descriptive capacity and are less effective for hypothesis generation (Evans et al. 2015). Statistical models are dependent on the specific input data (Austin 2007) thus

reducing their ability to make predictions using different data. Mechanistic models, on the other hand, provide an increased ability to test and explore the consequences of biological or environmental change (Kendal 1999) providing a more complete understanding of the dynamics of the modelled system. However, they generally require a larger amount of system-specific data (Nathan et al. 2001) or reasonable estimates for all modelled processes. As a specific example of these trade-offs, Kearney and Porter (2009) state that while statistical species distribution models are simpler to parameterize and often more generally applicable across species, mechanistic species distribution models may be more useful in some contexts if their data needs can be properly satisfied. They highlight specific research examples within the contexts of management strategy development, modelling of invasive species and more, suggesting that models emphasizing mechanistic processes were more effective than correlative techniques. Finally, Evans et al. (2015) argue that achieving generality in models primarily through simplicity provides only a superficial effect as the generality produced illustrates only possible outcomes but cannot explain specific outcomes. They suggest that complex models can also lead to general theory through simulation experiments conducted with variable parameters and different data sets. Also, through systematic parameterization and testing, processes and parameters can be organized in terms of importance of influence ecological patterns.

With this summary of mechanistic modelling in mind, I add two more points to the previous list of modelling objectives<sup>1</sup>:

*Identification of knowledge gaps:* assuming all existing knowledge of a system has been incorporated into a model, lack of consensus or poor fit between observed data and simulated results may indicate a gap in current knowledge. This gap identification can help to direct empirical study.

To illustrate this point, I consider the steps involved in developing a mechanistic model. Initial development begins with the modellers' own knowledge of the system and processes. Literature review will often provide parameter estimates and alternate approximations of processes but will also highlight gaps where data is not yet available. These gaps may exist where models are not used to guide empirical study and so the need for such data is overlooked. An example from the current work

---

<sup>1</sup>Similar to Baumgärtner et al. (2008) we do not strive to present an exhaustive list but rather to illustrate the breadth of objectives for which modelling is often undertaken. For more information see Baumgärtner et al. (2008), Grimm and Railsback (2003), Duarte et al. (2003), Hilborn and Mangel (1997), Gillman (2009), Mayer (1989).

is the lack of data on extrinsic (predation, disease, etc.) mortality rates of *D. suzukii* (see Chapter 2). While this is a significant consideration in the population predictions no attempts have yet been made to measure these rates.

*Organization of information:* presentation (often graphical or textual) and summary models can help organize complex information sets. Also, the process of developing conceptual or applied models can help the modeller organize and frame information related to a problem or system.

Similar to the previous point on revealing data gaps, mechanistic models can help to organize information. As all of the biologically relevant processes are constructed from research and literature, properly refuting the model results requires an organized understanding of that literature.

## 1.4 Modelling Invasive Insect Pests

Perhaps due to their relatively simple physiology and the availability of laboratory data there appears to be a growing number of insect-related mechanistic models. Some of the models are general explorations of behaviour (Beshers and Fewell 2001, Davis and Pelsor 2003) but the majority are concerned with economic impacts of invasive species management and commercial crop damage potential (Liebhold and Tobin 2008, Newman 2003, Kendall et al. 2005, Johnson 2005). The need for mitigation and management strategies is clear considering crop losses estimated over \$13 billion (Pimentel et al. 2005) due to invasive pest species alone. Over \$500 million dollars in pesticides are applied each year just to mitigate non-native insect pests. In addition to the financial cost, such measures have an environmental impact on native insects, plants and animals as well as humans and need to be considered. Models of population dynamics and distribution potential can play an essential role in developing management strategies by helping to target and justify costly treatments (Guisan et al. 2013, Chornesky et al. 2005) where they will likely be most effective. For these reasons models are considered an import component of integrated pest management (IPM) strategies (Thomas 1999, Way and Van Emdem 2000).

While many models are specifically concerned with the immediate impact of pest species, modelling shifts in species distribution due to climatic change also appears to be a growing area of

research interest. As temperatures increase, cooler regions that have historically been less susceptible to invasion by cold-intolerant species may become more habitable (Parmesan and Yohe 2003, Bale et al. 2002). Conversely, regions that are currently suitable may become too warm, exceeding thermal tolerance, and infestation risks may decline (Parmesan 2006). While there remains significant uncertainty in the specifics of how temperature (and other environmental factors such as humidity and precipitation) will change in the coming decades, existing data produced by General Circulation Models (GCMs) provide a potential range. Modelers and researchers are already using these data to provide preliminary risk assessment potential for various species (Berzitis et al. 2014, Crozier and Dwyer 2006, Newman 2006, Bentz et al. 2010, Kingsolver et al. 2011, Ziter et al. 2012). These efforts are part of proactive, long term management strategies (Estay 2009).

## 1.5 Modelling *D. suzukii*

Despite its extensive range, *D. suzukii* is a relatively novel pest in Europe, South America and North America. As a result, there are very few population dynamics models at this time despite efforts to establish a specific IPM strategy (Dreves 2011, Diepenbrock 2016). Among the first to be reported are a CLIMEX model of species distribution and a MaxEnt model of climate suitability (Damus 2009). CLIMEX attempts to predict potential distribution based solely on data from current existing location data (i.e., species-specific biological mechanisms are not considered) while MaxEnt can be used to estimate climate similarity between locations (i.e., biological response to climate is not considered). Damus (2009) found that the CLIMEX model produced a larger potential range for *D. suzukii* than what the MaxEnt climate suitability suggested but hypothesized that poor parameterization in CLIMEX may account for some of the discrepancy. Another early modelling effort is the *D. suzukii* sub model that is included in a larger online multi-pest phenology model (Coop 2010). This model highlights degree days based on selectable temperature profiles and reports results including estimated adult emergence and new generation dates. To date, no published (or unpublished) information on the model assumptions or design have been found.

In a similar approach, Wiman et al. (2014) developed a discrete time, phenological model based on the empirical life stage data of Tochen et al. (2014). Their model uses Leslie matrices to estimate the distribution of adults, eggs, pupa and larva for various temperature profiles. In general, Leslie matrices are used to study age-structured population dynamics and can be represented using

standard mathematical notation. The explicit age structure can be useful during sensitivity analysis which can highlight vital rates that are most influential. However, as with other mechanistic models, Leslie matrices have relatively high data requirements and in some cases, may require simplifying assumptions such as constant environmental conditions or constant vital rates. Validation attempts of the Wiman et al. model were largely unsuccessful as observed trap data did not correlate well with simulated results though the authors infer that limitations and concerns with trap data need to be addressed. They conclude that the model's utility lies not in absolute quantification but in relative projection of infestation potential and crop pressure. Using a more mechanistic approach, Gutierrez et al. (2016) developed a model of *D. suzukii* using empirical data from various sources including Tochen et al. (2014), Ryan et al. (2016), Emiljanowicz et al. (2014), and Dalton et al. (2011). Unlike the Wiman et al. model, the Gutierrez model aggregates the vital processes of development, fecundity and mortality for all life stages into a single egg-to-adult super-stage and represents time continuously.

In chapter 2, I present a significantly more mechanistically-rich (see Chapter 4) model of *D. suzukii* population dynamics. Similar to the models of Wiman and Gutierrez, the model processes and parameters are based on empirical laboratory observation of temperature-driven reproduction, mortality and development (Ryan et al. 2016, Emiljanowicz et al. 2014). Unlike both of the previously discussed models the one presented here includes considerations for all life stages including three larval instars and subdivided female stages to explicitly account for decreasing egg viability with age. Modular sub models include reproductive diapause generalized from literature on other *Drosophila* species and a simplified fruit quality index based on a degree-day model of sour cherries. Though somewhat more complex in design and presentation the model represents a more complete first-pass approximation of the current knowledge of *D. suzukii* population dynamics (with the exception of emerging research on humidity tolerance, cold-hardening and over-wintering). In addition to illustrating the base-model mechanics, I present the results of specific parameterization with observed temperature profiles for several important berry-producing regions across Canada and the United States. I show that the relative risk of infestation by *D. suzukii* is likely far from uniform across these locations suggesting that regionalized mitigation and management strategies may be more effective.

In chapter 3, I explore the potential for *D. suzukii* populations to shift and change over the coming decades. To do this I parameterize the model with temperature data derived from the CMIP5 family

of climate change models. Consensus maps of simulated populations are created for most of North America and show that in the absence of adaptation to increasing temperatures we may currently be at or near peak infestation potential for many regions.

In chapter 4, I present a more technical overview of the base model. A modified Overview, Design and Details (ODD, Grimm et al. 2006, 2010) protocol is followed to ensure that any users wishing to use, extend or modify the model have adequate information to do so. In addition to the ODD requirements, a discussion on parsimony as it relates to the population model and a usage guide are also presented.

## 2. North American Invasion of Spotted-Wing Drosophila (*Drosophila suzukii*): A Mechanistic Model of Population Dynamics<sup>2</sup>

### 2.1 Abstract

*Drosophila suzukii* is a relatively new threat to the soft-skinned fruit industry in North America. The presence of this pest in North America is a concern and assessing the risk of potential infestation and damage can help guide regional management strategies. I have developed a continuous time, stage structured, population model parameterized with empirical data based on laboratory observations. The principle environmental driver of vital rates (mortality, fecundity and development) for the model is temperature though my results suggest that reproductive diapause and quality of fruit available to the population may also have significant effect on population size. The model was run with several generalized temperature profiles and various observed temperature data sets for locations known to be important for berry production. While southern regions appear to be most suitable for supporting high population densities due to warm temperatures throughout the year, northern regions with moderate temperatures may also be susceptible due to a lack of extreme cold or heat, both of which limit development and increase mortality.

Keywords: *Drosophila suzukii*, mechanistic model, continuous time, temperature dependent development, reproductive diapause, fruit quality

---

<sup>2</sup>Published in part as: Langille AB, Arteca EM, Ryan GD, Emiljanowicz LM, Newman JA. North American invasion of Spotted-Wing Drosophila (*Drosophila suzukii*): A mechanistic model of population dynamics. Ecological Modelling. 2016 Sep 24;336:70-81.

## 2.2 Introduction

Reports of *Drosophila suzukii*, commonly known as Spotted Wing Drosophila (SWD), originating from the mainland of Japan have been present since the 1930s (Kanzawa, 1939). However, its rise as a pest of global significance has taken place largely over the past 35 years and in continental North America since circa 2009 (Hauser, 2011). *D. suzukii* has been confirmed or suspected present in countries across all continents except Australia and Antarctica (Asplen et al. 2015) and has a wide range of reported hosts including strawberries, raspberries, blueberries, blackberries, peaches, nectarines, pears, sweet and sour cherries, plums, apricots and both table and wine varieties of grape (Walsh et al. 2011).

*D. suzukii* is one of only two known species of Drosophila that prefer fresh, soft-skinned fruit at or near the stage of optimal harvest ripeness (Lee et al. 2011, Walsh et al. 2011). Females of this species possess a serrated ovipositor that allows them to attack pre-harvest fruit. Early detection of infestation is difficult as the oviposition incision is small and the visible consequences of larval feeding do not immediately appear. Empirical data for North America are not yet readily available but Bolda (2010) provides a benchmark estimate of potential yield losses due to *D. suzukii* infestation at 20%. Based on 2008 harvest yields, this would result in a >\$500 million loss in the US states of Washington, Oregon and California alone (accounting for strawberries, blueberries, raspberries, blackberries and cherries only). While the western United States and Canada are important regions for the soft-skinned fruit industry in North America there are other regions that are likely to be concerned about the spread and impact of *D. suzukii*.

Due to *D. suzukii*'s potential to cause significant economic losses in many areas of North America, and the need to establish effective pest management strategies, I developed a mechanistic, continuous time mathematical model with overlapping generations. Wherever possible, the model has been parameterized using experimental and observational data on *D. suzukii* life history (see Emiljanowicz et al. 2014, Ryan et al. 2016, for further information).

I have two principle objectives in constructing this model: (i) to gain a better understanding of the temporal and spatial patterns of SWD population growth, and hence risk, that result from particular assumptions about crucial mechanisms; and (ii) to create a useful tool for answering interesting “what if” questions regarding SWD invasion, climate and climatic change. A mechanistic model can help us to meet both of these objectives. Mechanisms are a critical part of biological understanding. “What if” questions usually involve changes in the environment,

autecology, or synecology of the organism. These questions require changes to the model system that are not readily accommodated in non-mechanistic approaches.

By mechanistic, I mean that the model is based upon what I think are the most important underlying processes that determine population dynamics. Such models are particularly useful because they allow any model prediction to be traced back to the process(es) that most influences it. Since mechanistic models often have many parameters, all of which are tunable, the goal is only to achieve satisfactory quantitative behavior. The goal is not forecasting, for which statistical models are often better suited (see e.g. Thornley and France, 2007, for further discussion).

### 2.2.1 *D. suzukii* Life History

*D. suzukii* develop through the following stages (with average duration reported in days based on laboratory diet and at a constant temperature of 22 °C): egg (1.4 d), three larval instars (6 d total), pupae (5.8 d), and adult (72 d) (see Emiljanowicz et al. 2014, for details).

Various *Drosophila* species are subject to a reproductive diapause regulated primarily by photoperiod and temperature (Ohtsu et al. 1993). As the number of hours of daylight decreases, reproduction ceases for several species (Saunders 1989, Kimura 1990). According to Saunders et al. (1989) diapause may be terminated by either an increase in the number of daylight hours, or more importantly for temperate regions, an increase in temperatures (approximately 18°C for *D. melanogaster*). It has been suggested that *D. suzukii* also experiences a reproductive diapause. Mitsui et al. (2010) found that females in populations of *D. suzukii* in central Japan had reproductively mature ovaries during summer months, while ovaries were underdeveloped for females trapped in colder months. However, to the best of my knowledge, the specific details of diapause induction and termination have not been elucidated for this species.

*D. suzukii* experiences both intrinsic and extrinsic mortality. Intrinsic mortality rate is the per capita rate of mortality assumed to be the result of biological aging. It can be contrasted with extrinsic mortality which is assumed to be the result of environmental hazards such as natural enemies (predators, parasites and pathogens). Intrinsic mortality rates have been estimated in the lab (Emiljanowicz et al. 2014) under optimal conditions. These rates are not likely to be density dependent, at least in the field; but field-based rates are difficult to estimate. Similarly, extrinsic

mortality rates are exceedingly difficult to estimate, at least in the field, which are arguably the only meaningful estimates of these rates.

*D. suzukii* vital rates are known to be temperature dependent, and development rates may also depend upon host fruit quality (Burrack et al. 2013, Lee et al. 2015). In the laboratory, optimal development rates occur at 28.2°C (Ryan et al. 2016; Tochen et al, 2014). Optimal fecundity rates occur at 22.9°C (Ryan et al. 2016; Tochen et al. 2014). The lower and upper developmental thresholds occur at 8°C and 31°C (Ryan et al. 2016) and 7.2°C and 42.1°C (Tochen et al. 2014). While the results are similar for most observations, I used the estimates provided by Ryan et al. because they were based on experiments that used a finer temperature scale, which, in particular, captured the upper developmental threshold rather than extrapolating it from the fitted curve as Tochen et al. did.

In terms of winter survival, Dalton et al. (2011) conducted laboratory experiments and found that acclimated adult *D. suzukii* can survive for up to 88 days in 10°C temperatures including a seven-day freeze period (from day 18 to day 25) but that adult longevity decreases as temperature decreases. Pupae can survive for 103 days at 10°C including the seven-day freeze period. Their results also suggest that *D. suzukii* is rendered sterile at these low temperatures. They conclude that few individuals are likely to survive the relatively moderate winters of the Pacific Northwest. More recently, Jakobs et al. (2015) found that 80% of control specimens died after 1 hour of exposure to cold temperatures (-7°C) and similar mortality rates were observed after a few days at 0°C (70–90h). Survivorship improved with cold acclimation and under fluctuating temperatures but, similar to the work of Dalton (2011) they conclude that *D. suzukii* is not well-adapted to survive temperate winters based on phenotypic characteristics alone. These results suggest that the presence of *D. suzukii* in climates with longer, colder winters is likely due to a combination of overwintering in man-made habitats, seasonal dispersal from warmer climates and developmental plasticity (but see Stephens et al. (2015) on the presence of winter-morphs). Due to the uncertainty that remains in how *D. suzukii* is overwintering in specific, particularly harsher climates, I do not model beyond a calendar year. As further details become available the model could be extended to include specific overwintering consideration.

## 2.3 Model

Population dynamics are represented by a system of coupled linear differential equations. The model presented here is a mechanistic population dynamics model based on the life stages of *D. suzukii*. There are 13 state variables: eggs, three juvenile instars, pupae, adult males and adult females and 121 parameters (see Appendix 2). The adult female population is divided into seven separate sub-stages in order to account for an observed decline in egg viability rates with age. The principal driving variable in the model is temperature as it is demonstrably related to the main population growth processes of fecundity, development and mortality. Wherever possible, parameters have been tuned and equations fit to approximate observed laboratory population dynamics, exceptions to this are noted in the text.

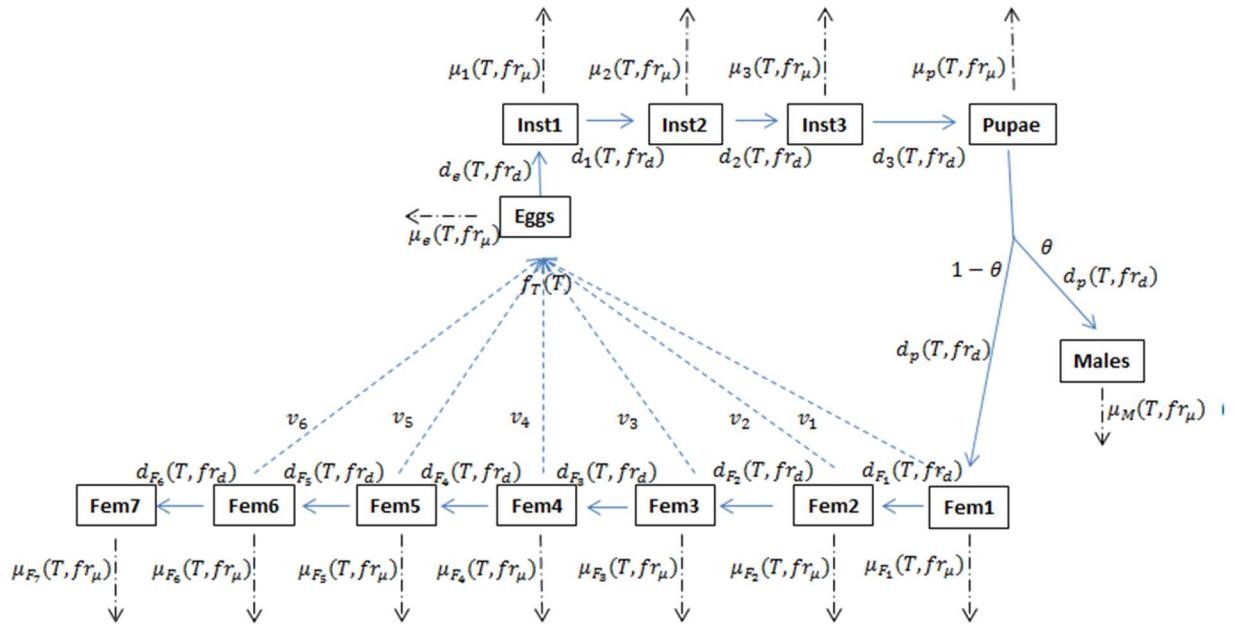


Figure 2.1 Schematic of the modelled population stage structure and relationships between state variable equations including principal input and output parameters. Most development and mortality processes are temperature ( $T$ ) dependent and rely on the fruit sub model ( $fr$ ) when included. Here mortality is represented as the sum of intrinsic and extrinsic factors (see main text).

### 2.3.1 Fecundity Rate

Laboratory experiments show that fecundity is dependent on temperature. I used a polynomial function with compact support (i.e., the function is both closed and bounded), similar to that described by Saryazdi and Cheriet (2007), to fit to the experimental data. This function is Gaussian-like, but unlike a Gaussian function, which is defined on the interval  $[-\infty, +\infty]$ , this function is constrained over a fixed interval  $[10, 30]$  where  $10^\circ\text{C}$  and  $30^\circ\text{C}$  were the lowest and highest temperatures respectively where fecundity was observed to be non-zero (Ryan et al. 2016).

$$f_T = \begin{cases} \alpha \left[ \frac{\gamma+1}{\pi \lambda^{2\gamma+2}} (\lambda^2 - ([T - \tau]^2 + \delta^2))^\gamma \right] & \text{if } T^2 + \delta^2 < \lambda^2 \text{ and } T_{f,\min} < T < T_{f,\max} \\ 0 & \text{otherwise} \end{cases} \quad \text{eqn 1}$$

where  $f_T$  is per capita eggs/day,  $T_{f,\max} = 30^\circ\text{C}$ ,  $\lambda = 52.68^\circ\text{C}$ ,  $\delta = 5.88^\circ\text{C}$  and  $\tau = 23.26^\circ\text{C}$  and the dimensionless shape parameters are given by:  $\alpha = 676.0$ ,  $\gamma = 88.38$  (see supplemental Figure A1.1; cf. Saryazdi and Cheriet for details regarding the constraint conditions on this function). It should be noted that the data used in the fecundity equations and parameters did not distinguish between female and male sterility.

I include the potential for a reproductive diapause through eqn 2 and eqn 3. Diapause data from Kimura (1990) including *Drosophila auraria*, *D. biauraria*, *D. subauraria*, and two geographic strains, *D. triauraria OI* and *D. triauraria ON*, was used to fit a generalized logistic function whereby  $F_D(h)$  provides the proportion of females currently in reproductive diapause based on photoperiod in hours,  $h$  at the current time step:

$$F_D(h) = \left( A + \frac{K-A}{(1+Qe^{Bh})^v} \right) \times 100^{-1} \quad \text{eqn 2}$$

with dimensionless shape parameters fit to  $A = 0.04$ ,  $K = 99.8$ ,  $v = 0.813$ ,  $Q = 3.23 \times 10^{-16}$  and  $B = 2.87$  per hour.

The following equations ensure that diapause is terminated by an upper critical temperature ( $T_D$ ) and is induced according to the number of daylight hours ( $h_D$ ).

$$s_{1,t+dt} = \begin{cases} 0 & \text{if } s_{1,t} \cdot s_{2,t} > 0 \text{ and } h < h_D \\ 1 & \text{if } s_{2,t} = 0 \text{ and } T > T_D \\ s_{1,t} & \text{otherwise} \end{cases}$$

$$s_{2,t+dt} = \begin{cases} 0 & \text{if } s_{1,t} = 0 \\ 1 & \text{if } h \geq h_D \\ s_{2,t} & \text{otherwise} \end{cases}$$
eqn 3

where  $s_1$  and  $s_2$  have no biological meaning; rather they act as computational control “switches” to enable and disable diapause, where  $h_D = 10$  (hours of daylight) and  $T_D = 18$  °C. I consider both values to be conservative estimates based on a review of the literature on diapause induction and termination of various Drosophilid species. Equations 2 and 3 are combined to determine a diapause-related fecundity factor:

$$f_D(h, T, t) = \begin{cases} 0 & \text{if } s_1 = 0 \\ 1 - F_D(h) & \text{otherwise} \end{cases}$$
eqn 4

Eqn (4) returns the proportion of females *not in* reproductive diapause. Overall fecundity (per capita eggs/day) is calculated using equations 1 and 4 as:

$$f = f_T f_D$$
eqn 5

### 2.3.2 Juvenile Development Rate

In order to express the temperature-dependent development of the egg-to-pupal stage, a function similar to one presented in Briere et al. (1999) was fit to experimental laboratory data (Ryan et al. 2016).

$$\ddot{d}_i = m_i \left( aT(T - T_L)(T_U - T)^{\frac{1}{2}} \right)$$
eqn 6

where  $a = 0.0001113$  is a dimensionless shape parameter, and where  $T$  is the air temperature (°C),  $T_L = 8.0139$  °C is the lower temperature threshold, and  $T_U = 30.99$  °C is the upper temperature threshold.

The multipliers  $m_i$  are the reciprocal of the number of days (days<sup>-1</sup>) required to develop from one stage to the next ( $m_e = 0.104$ ,  $m_1 = 0.0.082$ ,  $m_2 = 0.112$ ,  $m_3 = 0.231$ ,  $m_p = 0.470$  where the subscripts denote: eggs, instar 1, instar 2, instar 3 and pupae, respectively) and are used to scale the stage specific development accordingly. As experimental data were available only for the combined egg-to-pupal stage, I used the stage-specific development time in constant (optimal) temperature to determine the fraction of time spent in each stage. I then applied those fractions to the temperature dependent egg-to-pupal data to estimate the time spent in the intermediate stages.

### 2.3.3 Temperature-Dependent Mortality

The state variable equations include daily per capita mortality estimated from intrinsic and extrinsic causes. The equation below is minimized at optimal temperature and is an appropriate shape to approximate mortality from intrinsic causes including temperature-induced mortality.

$$\mu_i = \begin{cases} \sum_{j=0}^3 \beta_{i,j}(T - \tau)^j & \text{if } T_{\min,\mu} \leq T \leq T_{\max,\mu} \\ \mu_{i,\max} & \text{otherwise} \end{cases} \quad \text{eqn 7}$$

where  $i$  denotes the life stage and  $j$  is an exponent ranging from 0 to 3 representing the linear, quadratic and cubic effects of temperature,  $\beta_{i,1} = -0.008$ ,  $\beta_{i,2} = 0.00032$ ,  $\beta_{i,3} = -0.000002$ ,  $\tau = 8.178$ ,  $T_{\min,\mu} = 3.0$  °C and  $T_{\max,\mu} = 33.0$  °C.  $\tau$  is a horizontal shift and  $\beta_{i,0}$  is a vertical shift of the polynomial.  $T_{\min,\mu}$  and  $T_{\max,\mu}$  are the temperature thresholds beyond which mortality is maximized and  $\mu_{i,\max}$  represents the maximum per capita daily intrinsic mortality.  $\beta_{i,0}$  and  $\mu_{i,\max}$  are stage-specific parameters estimated using non-linear regression from Emiljanowicz et al. (2014) and are listed in Table 2.1.

$i$	Eggs	Instar1	Instar2	Instar3	Pupae	Males	F1	F2	F3	F4	F5	F6	F7
$\beta_{i,0}$	0.160	0.140	0.085	0.086	0.061	0.134	0.069	0.090	0.201	0.051	0.301	0.251	0.330
$\mu_{i,max}$	0.328	0.269	0.102	0.107	0.030	0.097	0.054	0.120	0.450	0.0	0.750	0.600	0.837

Table 2.1 Estimated parameters that vary by stage for temperature-dependent mortality (see e.g. Emiljanowicz et al. 2014). The  $\beta_{i,0}$  parameters are dimensionless shape parameters, and the  $\mu_{i,max}$  are per capita daily mortality rates.

In the absence of an empirical estimate for the  $\mu_{i,max}$  values, I assumed that these were equal to three times their corresponding values estimated at constant optimal temperature. I evaluated the sensitivity of the model results to this assumption and found that the results simply scaled to this assumption, but that the resulting dynamics were largely unaffected (see supplemental Figure A1.2).

### 2.3.4 Effect of Fruit Quality on Development and Mortality

To this point, I have assumed that fruit is always available at a quality that maximizes *D. suzukii* development. In order to explore how seasonal fruit availability affects the seasonal population dynamics of *D. suzukii* I implemented a non-mechanistic model of fruit that captures the general phenomenon and temporal dynamics of the ripening and harvesting of a generalized fruit crop. A differential equation was constructed to simulate this process:

$$\frac{dFr}{dt} = \left( \frac{\omega}{G_{Fr}(T)} - D_{Fr} \right) Fr \quad \text{eqn 8}$$

$$Fr = \begin{cases} 0.05 & \text{if } Fr < 0.05 \\ 1 & \text{if } Fr > 1 \\ Fr & \text{otherwise} \end{cases}$$

where  $\omega$  is a dimensionless shape parameter that determines when fruit will be at optimal ripeness and  $Fr$  is a dimensionless index ranging from 0.05 to 1 representing the current quality of fruit present (see Figure 2.2a).  $G_{Fr}$  (eqn 10) is the fruit quality development time and  $D_{Fr}$  (eqn 9) is fruit quality decline rate. Previous studies have shown that *D. suzukii* can subsist on various wild hosts when cultivated fruit is unavailable (Lee et al. 2015, Walsh et al., 2011). As such, I assume that non-preferred

fruit is always available during the simulation, albeit at a minimized quality level, by including a lower bound of 0.05 on  $Fr$ . Fruit quality  $D_{Fr}$  declines after harvest time lag has elapsed and is calculated as:

$$D_{Fr}(t) = \begin{cases} 0 & \text{if } Fr(t - h_{Fr}) < \varphi \\ \gamma & \text{otherwise} \end{cases}, \quad \text{eqn 9}$$

where  $t$  is time and  $\varphi$  is the maximum fruit quality value.  $h_{Fr}$  represents a harvest time lag, that is, the amount of time that passes between when the fruit is at optimal quality and when it is harvested and removed from the simulation.  $\gamma$  is the constant rate of fruit quality decline over time. At this time, fruit development is modeled to be a single-stage (i.e., not stage specific).

Temperature-based fruit development time,  $G_{Fr}$ , is based on a phenological degree-day growth model of sour cherries. The double sigmoid fruit growth model of Zavalloni et al. (2006) was used to calculate the number of days required at various temperatures for sour cherries to reach optimal fruit ripeness. These data points were then fit to the following function:

$$G_{Fr}(T) = \frac{1100}{T - T_{base}} + 30 \quad \text{eqn 10}$$

where  $T_{base}$  is a vertical asymptote at 4.0°C representing the base temperature required for the fruit to develop. As temperature increases the time required for the fruit to develop,  $G_{Fr}$ , decreases.

I make the simplifying assumption that fruit is always available to the population and that only the quality changes with time. Fruit below optimal host quality decreases the rate of development of the simulated population. I constructed the following index of fruit quality (eqn 11) that is used to modify *D. suzukii* development rates (*cf.* Newman et al., 2003). This equation is mathematically transparent and of the correct general shape, acting as a “switch on” sigmoid controlling the effect of fruit quality on development.

$$fr_d = m_{fd} \left( \frac{Fr}{Q_{p,fd,h}} \right)^{n_{fd}} \left( 1 + \left( \frac{Fr}{Q_{p,fd,h}} \right)^{n_{fd}} \right)^{-1} + 1 - m_{fd}; \quad 0 \leq m_{fd} \leq 1 \quad \text{eqn 11}$$

where,  $Fr$  (eqn 8) is the simulated current host fruit quality value and  $m_{fd} = 0.75$ ,  $Q_{p,fd,h} = 0.5$  and  $n_{fd} = 4$  are shape parameters;  $m_{fd}$  influences the strength of the fruit quality response,  $Q_{p,fd,h}$  is a half

saturation constant and  $n_{fd}$  controls the rate of switching (see Figure 2.2b). The index (eqn 11) ranges from  $1 - m_{fd}$  to 1, and is used to modify *D. suzukii* development (eqn 6) as:

$$d_i = \check{d}_i f r_d \quad \text{eqn 12}$$

Similarly, the effect of fruit quality on mortality is calculated using the following “switch-off” sigmoid:

$$f r_\mu = m_{f\mu} \left( 1 + \left( \frac{Fr}{Q_{p,f\mu}} \right)^{n_{f\mu}} \right)^{-1} \quad \text{eqn 13}$$

where  $m_{f\mu} = 0.1$ ,  $Q_{p,f\mu} = 0.5$  and  $n_{f\mu} = 4.0$  are shape parameters;  $m_{f\mu}$  influences the strength of the mortality response,  $Q_{p,f\mu}$  is a half saturation constant and  $n_{f\mu}$  controls the rate of switching (see Figure 2.2b). I consider the effect of fruit host quality on mortality to be additional to the mortality due to temperature. This is a phenomenological representation and I consider it to be equivalent across all life stages.  $f r_\mu$  creates an additional mortality that ranges from 0 when fruit quality is optimal to 1 when fruit quality is minimal. I thus model intrinsic mortality,  $\mu_{in}$  as:

$$\mu_{in} = \mu_i + f r_\mu \quad \text{eqn 14}$$

where  $\mu_i$  is given by eqn 7.

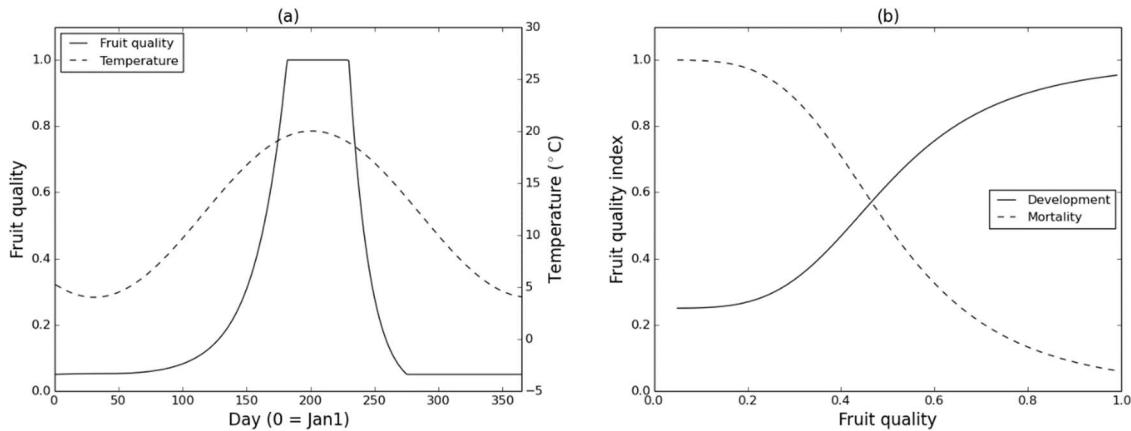


Figure 2.2 a) Fruit quality development time as described in eqn 8 for a simulated temperature profile (see section 2.6.2) with a harvest delay,  $\omega$ , 50. b) illustrates fruit quality index (eqn 11) for influencing development and fruit quality decline (eqn 13) for influencing mortality.

### 2.3.5 Model State Equations

The following equations are used to model the rates of change of the population of the various stages of *D. suzukii*'s life history. The rate of change for the egg stage is given by:

$$\frac{dE}{dt} = (\sum_{i=1}^6 fv_i F_i) - E(\mu_{in,e} + \mu_{ex,e} + d_e) \quad \text{eqn 15}$$

where  $f$  is fecundity (eqn 5),  $E$  is the number of eggs at time  $t$ , and  $v_i$  represents egg viability at a particular adult female ( $F_i$ ) stage. Data on female age-specific egg viability were obtained from Emiljanowicz et al. (2014). This is depicted in Figure 2.3 (see figure caption for details). Egg daily per capita intrinsic mortality,  $\mu_{in,e}$ , is given by eqn 14, and egg daily per capita extrinsic mortality,  $\mu_{ex,e}$ , represents the rate of egg loss due to natural enemies.  $d_e$  represents the daily per capita rate of eggs developing into first instar larvae ( $I_1$ , see eqn 16).

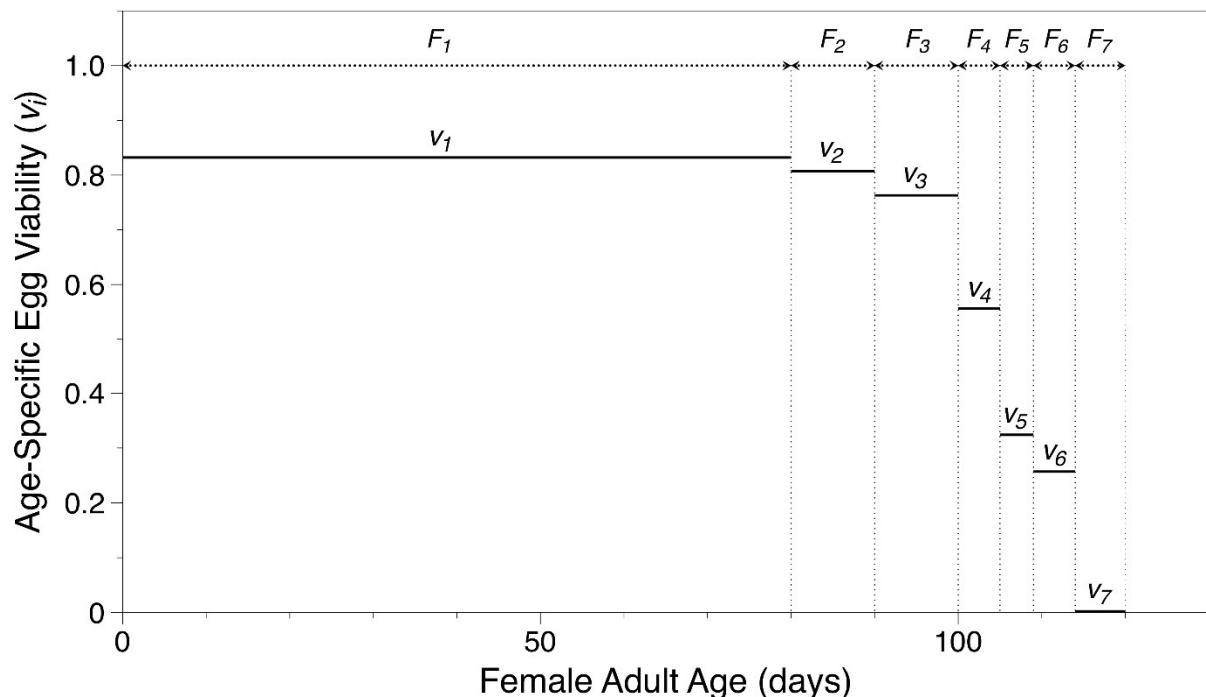


Figure 2.3 Female age-specific egg viability. Changes in viability correspond to  $v_i$  in eqn 15 where  $v_1 = 0.832$ ,  $v_2 = 0.807$ ,  $v_3 = 0.763$ ,  $v_4 = 0.556$ ,  $v_5 = 0.324$ ,  $v_6 = 0.257$  (see Emiljanowicz et al. 2014 for further information).

The rates of change for each of the three larval instars ( $I_1, I_2, I_3$ ) are given by:

$$\begin{aligned}\frac{dI_1}{dt} &= d_e E - I_1(\mu_{in,1} + \mu_{ex,1} + d_1) \\ \frac{dI_i}{dt} &= d_{i-1} I_{i-1} - I_i(\mu_{in,i} + \mu_{ex,i} + d_i) \quad \text{for } i = 2, 3\end{aligned}\tag{eqn 16}$$

again, where  $\mu_{in,i}$  and  $\mu_{ex,i}$  represent the daily per capita intrinsic and extrinsic mortality rates at each larval instar ( $i = 1, 2, 3$ ) and the  $d_i$  represent daily per capita development rates ( $i = 1, 2, 3, e$ ).

The rate of change of the pupal stage is as follows

$$\frac{dP}{dt} = d_3 I_3 - P(\mu_{in,p} + \mu_{ex,p} + d_p)\tag{eqn 17}$$

where  $d_3 I_3$  is the rate of individuals transitioning to pupae from the third instar and  $\mu_{in,p}$ ,  $\mu_{ex,p}$  and  $d_p$  represent the daily per capita lost to intrinsic mortality (eqn 14), extrinsic mortality and development to the adult stage (eqn 12), respectively.

The rate of change of adult males is given by

$$\frac{dM}{dt} = \theta d_p - M(\mu_{in,M} + \mu_{ex,M})\tag{eqn 18}$$

where  $\theta = 0.5$  and represents the proportion of adult flies that are males.

The adult female population is divided into seven sub stages, each with its own rate of change, to account for declining egg viability as females age (see Figure 2.3).

$$\begin{aligned}\frac{dF_1}{dt} &= (1 - \theta)d_p P - F_1(\mu_{in,F_1} + \mu_{ex,F_1} + d_{F_1}) \\ \frac{dF_i}{dt} &= d_{F_{i-1}} F_{i-1} - F_i(\mu_{in,F_i} + \mu_{ex,F_i} + d_{F_i}) \quad \text{for } i = 2 \dots 7\end{aligned}\tag{eqn 19}$$

Pupae develop into adult flies at a daily rate of  $d_p P$ . The proportion of pupae that become adult female flies ( $F_1$ ) is  $1 - \theta$ . Different female stages ( $F_i$ ) were derived from empirical data on female aging and reproductive output (Emiljanowicz et al. 2014) and reflect the female stages over which offspring viability was found to be approximately constant (see Figure 2.3). Daily per capita development rates for each of these stages were estimated from laboratory experiments (Emiljanowicz et al 2014):

$d_{F_i} = \frac{1}{80}, \frac{1}{10}, \frac{1}{10}, \frac{1}{5}, \frac{1}{4}, \frac{1}{5}, 0$  for  $i = 1$  to 7. The final adult female stage ( $F_7$ ) has no development rate as there is no stage beyond  $F_7$  but mortality continues per eqn 19, and they produce no viable eggs (Figure 2.3).

Although the model appears to be complex (Figure 2.1) it is conceptually simpler when one considers that both the three larval instars and seven adult female stages are replicate stages with similar function. While simpler aggregate versions of these stages might be used, I avoid doing so here in order to maintain the full mechanistic nature of the model thereby maximizing its elucidation and exploration potential (see section 4.2.4 for further discussion).

### 2.3.6 Environmental Equations

Equations related to diapause termination (eqn 2 and eqn 3) require number of daylight hours,  $h$ , which is calculated as follows (Glarner 2006):

$$h = \frac{24}{\pi} \times \arccos(k') \quad \text{eqn 20}$$

where  $k'$  is computed as

$$k' = \begin{cases} -1 & \text{if } k < 1 \\ 0 & \text{if } k > 1 \\ k & \text{otherwise} \end{cases} \quad \text{eqn 21}$$

Finally,  $k$ , the exposed radius between the sun's zenith and the sun's solar circle, is a function of latitude,  $L$ , the number of days from January 1<sup>st</sup>,  $\delta$ , and the Earth's rotational axis,  $\alpha_e = 23.439^\circ$ .

$$k = \tan\left(\frac{\pi L}{180^\circ}\right) \tan\left(\left(\frac{\pi R}{180^\circ}\right) \cos\left(\frac{\pi Y}{182.625^\circ}\right)\right) \quad \text{eqn 22}$$

In order to demonstrate the base model's behavior, I use a cosine model of daily mean temperature (eqn. 23). For the figures shown in the results section, I parameterized the cosine model to approximate temperature records for four distinct geographical regions across Canada and the United States. The mean 20-year (1993-2013) observed daily mean temperature data for Chicoutimi

township, Quebec; Clark county, Washington; Santa Barbara county, California; and Hillsborough county, Florida are reasonably represented by Temp A, B, C and D respectively in Figure 2.4 (see also Table 2.3). Each of these ‘profiles’ was generated using eqn 23, where  $\varepsilon$ ,  $k$ ,  $s$  and  $w$  represent amplitude ( $^{\circ}\text{C}$ ), horizontal stretch ( $\text{days}^{-1}$ ), horizontal shift ( $^{\circ}$  latitude) and vertical shift ( $^{\circ}\text{C}$ ) respectively, and are parameterized as per Table 2.2.

$$T = \varepsilon \times \cos \left( \frac{k\pi t}{365} - \frac{s\pi}{180} \right) + w \quad \text{eqn 23}$$

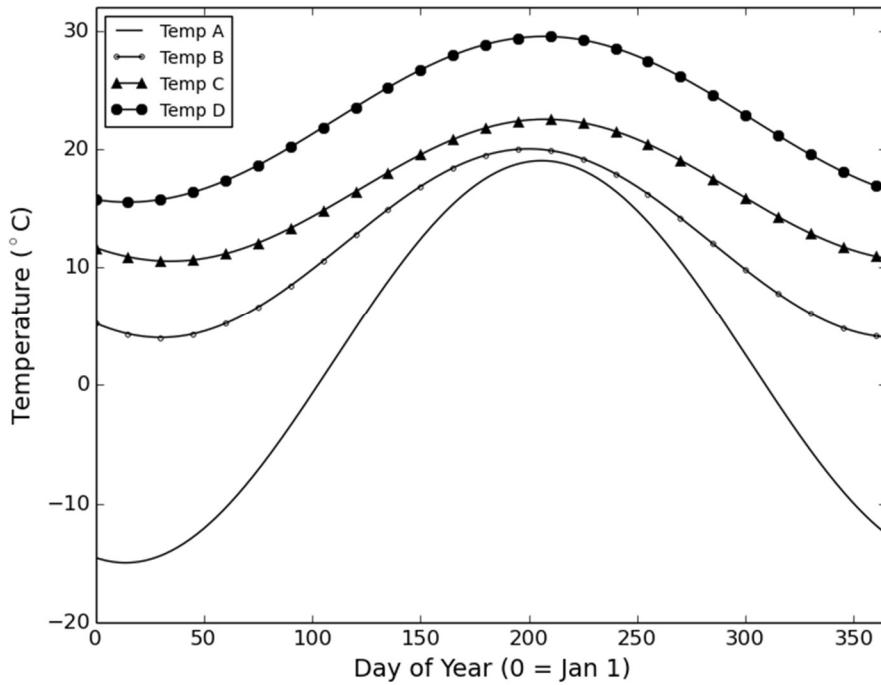


Figure 2.4 Modelled temperature profiles approximating 20-year mean daily temperature for Chicoutimi QC (A), Clark county WA (B), Santa Barbara county CA (C) and Hillsborough county FL, (D).

Profile Name	$\varepsilon$ ( $^{\circ}\text{C}$ )	$k$ ( $\text{days}^{-1}$ )	$s$ ( $^{\circ}\text{lat}$ )	$w$ ( $^{\circ}\text{C}$ )	Profile Temperature Characteristics
Temp A	-17	1.9	13	2	Moderate summer, cold winter, large inter seasonal variation
Temp B	-8	2.15	32	12	Moderate summer, cool winters, moderate inter seasonal variation
Temp C	-6	2.11	36	16.5	Warmer summer, moderate winter, reduced inter seasonal variation
Temp D	-7	1.9	14	22.5	Hot summers, moderate winter, reduced inter seasonal variation

Table 2.2 Parameters used to generate model temperature profiles per eqn 23.

Observed mean daily temperature data was collected for a 20-year span (1993-2013) for

various soft-skinned fruit producing counties and townships across the United States and Canada (see Table 2.3). U.S. data was obtained via the National Climatic Data Center (<http://www.ncdc.noaa.gov/>) and Canadian data was obtained via Environment Canada (<http://climate.weather.gc.ca/>).

County/Township, State/Province	Aproxmimate Latitude, Longitude
Fraser Valley, BC	49.5800° N, 121.8333° W
Chicoutami, QC	48.4200° N, 71.0500° W
Kent, NB	46.5800° N, 64.8000° W
Cumberland NS	45.7000° N, 64.1000° W
Clark, WA	45.7700° N, 122.4800° W
Norfolk, ON	42.8500° N, 80.2600° W
Allegan, MI	42.5600° N, 86.2500° W
Burlington, NJ	39.8800° N, 74.6700° W
Bladen, NC	34.6200° N, 78.5600° W
Santa Barbara, CA	34.5400° N, 120.0300° W
Wayne, MS	31.6400° N, 88.7000° W
Hillsboro, FL	27.9100° N, 82.3500° W

Table 2.3 Latitude and longitude of 12 of important blueberry and/or strawberry producing counties and townships in the United States and Canada. Sorted by latitude (north to south). See Figure A1.3 in the supplemental material for a map of these locations.

### 2.3.7 Model Implementation

The model was implemented and all of the equations presented in the previous sections were solved using C++, specifically GNU gcc version 4.8.4 for Ubuntu Linux (14.04.03 LTS). All differential equations (eqn 15 through eqn 19) were solved using Euler's method of numerical integration. Source code files including headers are freely available at <https://github.com/alangillGuelph/dsPopSim>.

## 2.4 Results

Unless otherwise stated all simulations begin with 10 females ( $F_1$ ) introduced during the time step for which temperature is adequate to break diapause ( $18^\circ\text{C}$ ) and are run for a full year starting January 1<sup>st</sup>. In all results, population counts are limited to females (as the limiting factor) and have been normalized by dividing all populations by the maximum population (per data set) such that the largest population is 1. This operation emphasizes relative rather than absolute differences and facilitates comparison across time and space. The per-simulation maxima used for normalization as well as a summary of sub model parameters are available supplemental material (Tables A1.1 through A1.3).

### 2.4.1 Base Model Dynamics

Figure 2.5 demonstrates the base population dynamics for a series of simulations where temperature was held constant. Total cumulative population (all adult female stages) over the entire simulation duration is reported and neither diapause nor fruit quality are considered initially. Provided temperatures are favourable for fecundity and development, population size increases exponentially. Cumulative population sizes increase until the optimal temperature is reached, after which development rates decrease, mortality increases and the simulated cumulative population declines. The resulting normalized cumulative population sizes under constant temperature are shown in Figure 2.5.

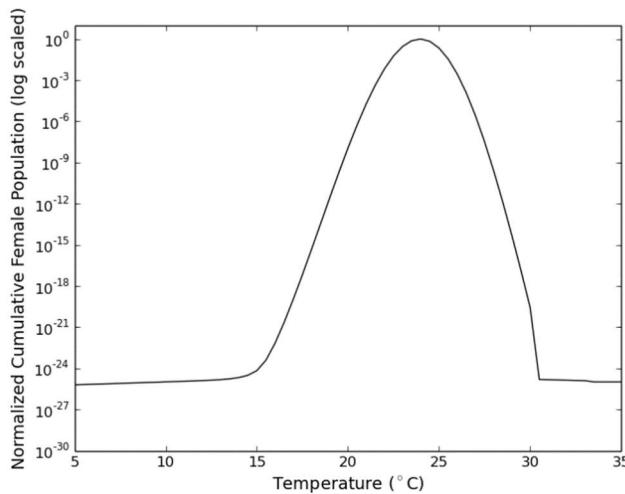


Figure 2.5 Normalized cumulative female population for various constant temperatures. Simulations are 365 time steps in length, begin with 10 fecund females and do not include the diapause or fruit quality sub models.

## 2.4.2 Effects of Starting Population Size and Appearance Timing

Since little is known about how *D. suzukii* overwinters and overwintering mortality rates can be decoupled from the actual air temperature by choice of microhabitat or use of heated human-made structures, I examined the sensitivity of the model results to the two principal variables that might vary with overwintering: the size of the population that survives the winter (or immigrates from a region with a warmer climate) and the date at which they appear in a region (Figure 2.6). This sensitivity test does not consider the diapause and fruit quality sub models introduced later. The initial population size was varied from 10 to 10000  $F_1$  females, and the day on which they are added to the simulation was varied from 0 (i.e. January 1<sup>st</sup>) through 364 (December 31<sup>st</sup>). Model temperature profile Temp B (see Table 2.2) is shown here for illustration. Figure 2.6 illustrates the resulting normalized cumulative population (all life stages) for different introduction dates and populations sizes.

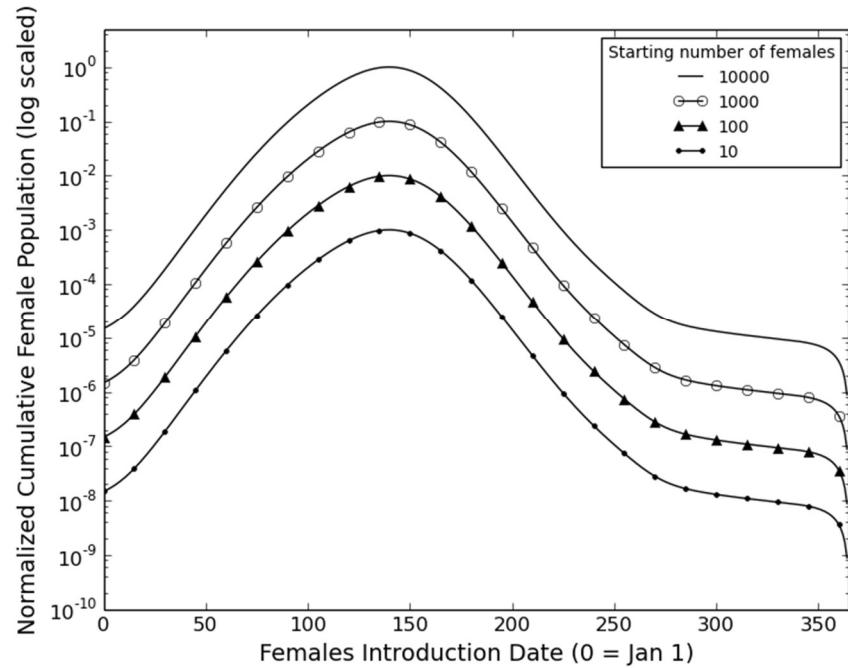


Figure 2.6 Cumulative population sizes for varied initial introduction dates and populations sizes. In both cases, introduced flies are fecund females and the simulated timeframe uses model temperature profile Temp B. Individual simulations were run for 365 time steps.

Due to the continuous population representation and a single introduction date, the time step at which the population is introduced has a strong impact on the cumulative population size. For very

early introduction dates the population can become extremely small despite having the maximum potential time for growth. Increased mortality as well as decreased development at colder temperatures produces a decline in the population below one individual and tending towards (but not reaching) zero during the colder first part of the simulated year. This creates a population ‘deficit’ that must then be overcome during the warmer periods. As the initial appearance day occurs later in the year this effect lessens and cumulative population increases significantly. In general this increase continues until the maximum cumulative population occurs with an introduction date that avoids decline due to early cold and allows for maximum time for the population to increase before temperatures decrease later in the year. In this simulation the optimal appearance day was May 21st. After this date the cumulative population decreases as declining temperatures in the fall and winter leave the population with less opportunity for positive growth. The optimal date of introduction is dependent on the location-specific yearly temperature profile, and is affected by the rate at which temperatures increase as well as the maximum temperatures which, when above heat tolerance, affect mortality and male sterility. Simulations run with temperature profiles that have less pronounced seasonality (regardless of peak temperature) are less sensitive to introduction date. Varying the starting population has a direct linear scaling effect on the cumulative population for a given introduction date.

#### 2.4.3 Diapause Sub Model

As stated in the introduction, the precise environmental control on diapause induction is not yet known for this species of *Drosophila*, but I made the reasonable assumption that it is similar to that found for other species of *Drosophila*. Sensitivity analysis of the diapause sub model suggests that the number of daylight hours required to induce diapause has little effect on the cumulative population size (see supplemental Figure A1.2). This is due to the fact that temperatures decline and population growth rates decrease before diapause induction actually occurs. Because of this minimal effect, the diapause induction parameter was fixed at a conservative 10 hours of daylight for all simulations.

Similarly, the precise temperature causing diapause termination is not yet known for this species. Based on work for other species of *Drosophila*, I estimated that 18°C was a reasonable first approximation for this parameter value. I investigated the impact of diapause termination

temperature on the cumulative population for all four model temperature profiles (Figure 2.7).

Unlike the number of daylight hours required for diapause induction, temperature-based diapause termination does have a large effect on the cumulative population. In particular, as the termination threshold temperature increases, the amount of time available for population growth is reduced, and thus the cumulative population size declines. It is worth noting that the effect of diapause is most noticeable in regions of moderate climate (i.e., Temp profile C). In regions where temperatures are warm and relatively constant throughout the year (Temp D), the population is less sensitive to this parameter, as the higher temperatures ensure that the population breaks diapause early and cumulative population size is maximized. Only when the diapause termination temperature is set to unrealistically high values does it affect the population. In cooler climates (Temp A and B) the effect of diapause termination is also minor as the temperatures do not reach the diapause threshold until well into the year and the temperature profile is not as favourable for population growth when diapause is broken.

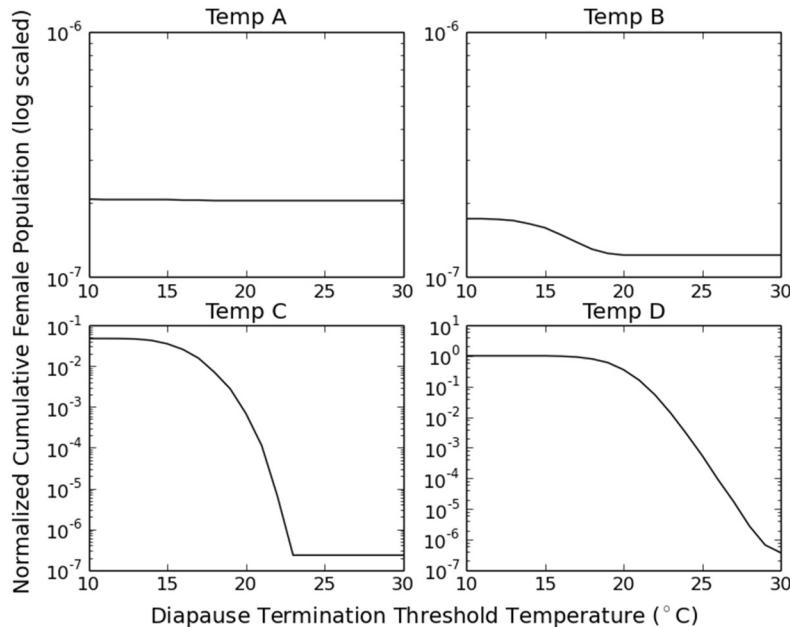


Figure 2.7 The effect of diapause termination temperature on cumulative populations for model temperature profiles Temp A through D. Each simulation starts with 10 fecund females introduced into the simulation at time step 0 and was run for 365 days. Diapause induction is fixed at 10 hours of daylight for all simulations. Note the difference in the y-axis scaling between graphs.

As discussed previously (Figure 2.6), the introduction of the initial flies when temperatures are too cool to promote positive development can cause a population deficit that may not be overcome, particularly when the diapause termination threshold is high relative to local warm season

temperatures. This population deficit is biologically unrealistic and in order to correct for this the initial individuals are added on the timestep when the temperature is adequate to break diapause. Figure 2.8 shows the combined effects of the base population dynamics and diapause sub model with various modelled temperature profiles.

Each of the illustrated populations shows similar growth followed by decline after peak temperature. Temp D has the earliest diapause termination date due to warmer temperatures earlier in the model year. The population increases quickly but the rate slows somewhat as temperatures exceed those required for optimal development. Temp C produces a later diapause termination date but surpasses the population size of Temp D as it remains at or near optimal development temperatures for a longer period. Temp A produces the lowest population peak due to the latest diapause termination date and temperatures consistently below optimal development.

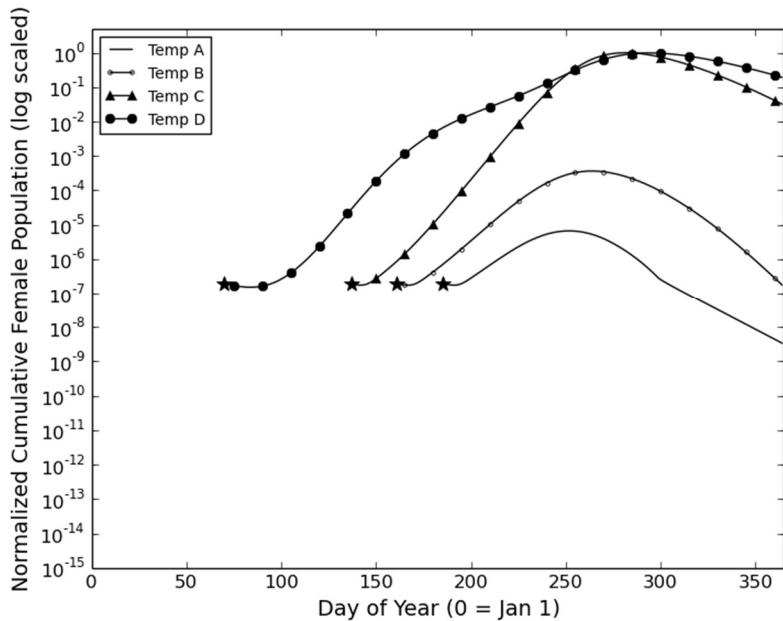


Figure 2.8 Normalized per-day female count for each of the model temperature profiles. 10 fecund females are introduced into the simulation on the time step (starred) for which diapause would be terminated due to adequate temperature.

#### 2.4.4 Observed Temperature Data simulations

In addition to the cosine temperature curves, I also used observed daily mean temperature records, and 20-year average daily temperature as a way of investigating the influence of inter-annual variability and climate. Figure 2.9 illustrates the population dynamics based on empirical temperature data. I conducted 21 separate simulations, 20 using observed daily average temperature for Clark County WA from 1993 to 2013 inclusive, and one using the 20-year averaged daily temperatures. The results show the normalized daily count of females (all stages).

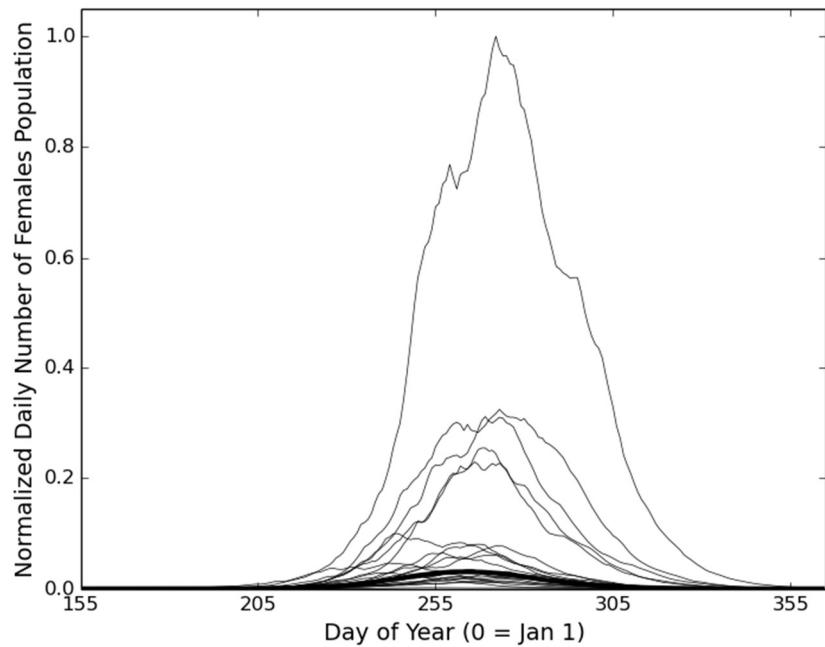


Figure 2.9 Normalized per-day female (all adult stages) population dynamics using observed daily mean temperature data for Clark County, WA (1993-2013). Bold line indicates simulation using 20-year average daily temperature. 10 females are introduced into the simulation on the time step where temperature is adequate to break diapause.

These results illustrate the flies' sensitivity to annual variation in temperature. I can look to several temperature cues to highlight the reasons for the large spread of values. The mean annual temperature for the year producing the largest population (2003) is 12.0°C (the second largest mean daily temperature of the 20-year data set) compared with 10.84°C for the year producing the smallest peak population (1996). In addition, diapause terminates on April 8<sup>th</sup> in 1996 and on May 23<sup>rd</sup> in 2003. While this would appear to give the 1996 population a 45-day growth advantage, instead

diapause terminates on an abnormally warm day early in the year and subsequently the population crashes when temperatures return to low seasonal norms. The later diapause termination date in 2003 ensures that the population begins and maintains a positive growth phase at a near-optimal time step for population growth. Comparing the 2003 population with that of 2004, which has the highest annual average daily temperature at  $12.14^{\circ}\text{C}$ , I note that 2004 has the fourth highest cumulative population, approximately one third that of 2003. Again, I find that an earlier diapause termination on April 27<sup>th</sup> results in a short phase of negative population growth followed by sub-optimal positive growth early in the simulation compared to the optimal growth of the 2003 population. Diapause termination date and mean daily temperature together contribute significantly to the cumulative population size.

#### 2.4.5 Fruit Sub Model

Up to this point, I have illustrated the behaviour of the model under the assumption that food is always available at an optimal quality for development. Figure 2.10 shows the effect on the cumulative population of varying the quality of available food through the fruit sub model. For these simulations 10 F<sub>1</sub> females were introduced on March 16th (day 75) rather than at diapause termination as the diapause sub model was not considered for these simulations. The  $\omega$  multiplier determines the rate at which fruit reaches optimal quality and the harvest lag determines the number of days in which optimal quality fruit will be available to the simulated population. The simulations were run for the temperature profile Temp B.

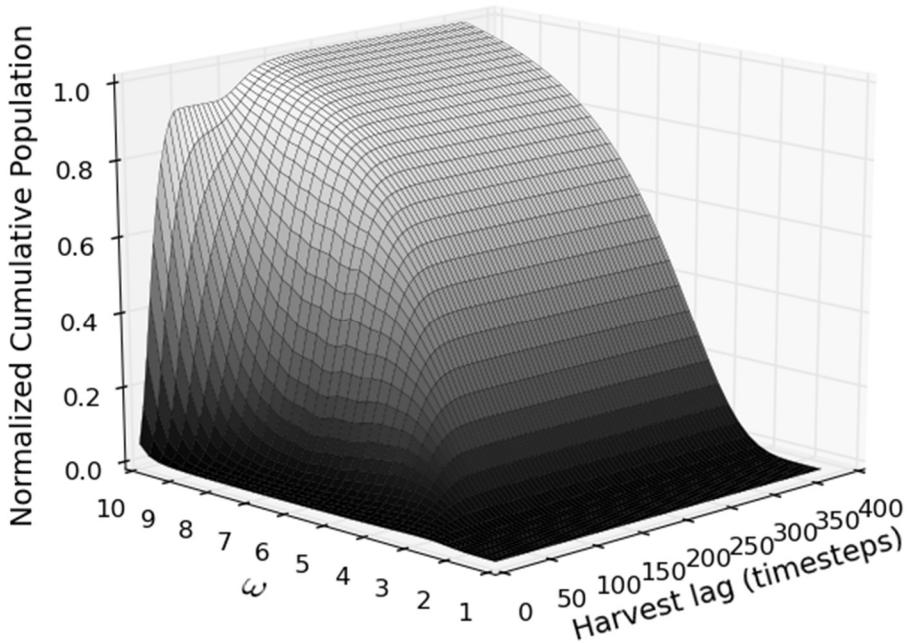


Figure 2.10 Fruit sub model effect on cumulative population. Simulations start with 10 fecund females introduced on time step 75 and are run for a duration of 365 time steps. Model temperature profile Temp B is used. Cumulative population is maximized as both  $\omega$  (the quality rate limiter) and harvest lag increase resulting in optimal food availability for the simulated population.

I note that as  $\omega$  increases so too does the population as fruit becomes available earlier in the year, thereby optimizing the opportunity for population growth. This effect is enhanced when the harvest lag increases, but only until the lag reaches approximately 100 days, after which factors such as temperature-driven mortality and late-year declining temperature and diapause initiation determine population thresholds. The effect of the fruit sub model is similar for various locations tested but it is worth noting that with a fixed  $\omega$  multiplier optimal fruit quality is reached sooner in regions with warmer temperatures while in regions with colder temperatures may not be reached at all preventing the population from reaching its maximum potential size. Furthermore, a short harvest lag time can inhibit population growth by limiting food source availability. This effect is most pronounced when  $\omega$  is very low (late ripening fruit) or very high (early ripening fruit) and when combined with a short harvest lag period.

## 2.4.6 The Complete Model

The complete model combines the effects of diapause and the fruit sub model. In Figure 2.11 I illustrate the complete model's behaviour. Normalized cumulative populations are shown for various strawberry and blueberry producing counties in the United States and townships in Canada. For each location daily average temperature data were obtained for the years 1993 through 2013. The simulation was run for 365 days and the mean normalized cumulative population, diapause termination date and date at which the peak number of adult females occurred were recorded for each year for each location. To illustrate the effects of the fruit quality sub model, the simulations are repeated with the same parameters and temperature data and the fruit parameters  $\omega$  set to 4.0 and harvest lag time set to 50 days. Also included are the dates where peak population (adult females) occurs for the model both including and excluding the fruit sub model.

In the simulation without the fruit quality sub model it is the southern-most counties that produce the largest simulated populations of *D. suzukii*. The relatively northern Burlington County, New Jersey appears to be an exception to this primarily due to warmer, but not extreme, temperatures throughout the year. The maximum temperature-based intrinsic mortality and upper threshold for temperature-based development are reached at 33°C and 31°C respectively, which explains why a region with more moderate temperatures such as Burlington county might have one of the larger populations. If the temperatures are near optimal for development (approximately 24°C) with minimal days above the threshold for maximum mortality or minimum development, then a population could grow optimally at a more northern latitude. The diapause termination date follows a largely south-to-north trend due to higher temperatures early in the year. However, there is little evidence of a significant latitudinal trend for peak population date due to a combination of declining temperatures and diapause induction in all locations later in the year.

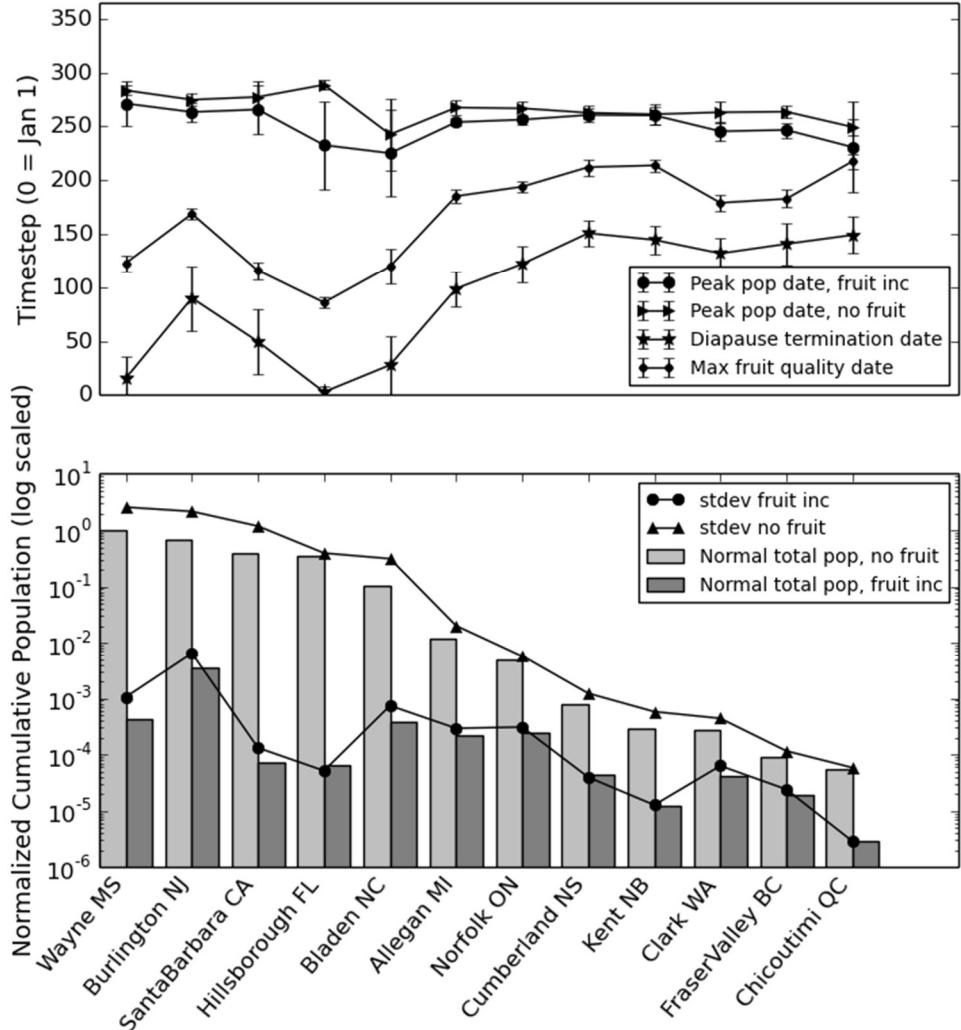


Figure 2.11 Population dynamics, including diapause and fruit quality considerations, for various soft-skin producing counties and townships across the United States and Canada. Cumulative population is normalized to show relative potential population sizes. Time step at which diapause would terminate and peak adult female population occurs for the diapause-only and diapause-fruit models are also shown.

Since the base model implies that food is available at optimal quality at all times smaller overall populations were observed when the fruit quality sub model was included (Fig. 9). Because development decreases and mortality increases when fruit is not at optimal ripeness, populations do not reach the maximums observed when the fruit sub model is not present. Peak population size is reached earlier in the year when the fruit sub model is included. This is the result of a decline in population after the harvest lag time has elapsed. When the fruit sub model is ignored populations can continue to grow provided diapause has not been initiated and temperatures are favourable. The

warmer regions in Hillsborough (FL), Santa Barbara (CA) and Wayne (MS) have lower cumulative populations than might be expected since the fruit in the model, also being temperature dependent, ripens very quickly. The population then grows optimally for the duration of the harvest lag time after which development rates decrease due to the limited availability of fruit for the remainder of the simulated year. On the other hand, regions where the fruit grows more slowly offer the population a sub-optimal quality yet available food source for a longer period of time followed by the same harvest lag time thus producing relatively larger populations.

#### 2.4.7 Extrinsic Mortality

Extrinsic mortality (i.e. mortality due predators, parasites and pathogens) in the field is inherently difficult to quantify but is likely a very important determinant of population size. The simulations shown in Figure 2.11 considered extrinsic mortality to be zero. Here I show the impact that increasing extrinsic mortality has on relative population size (Figure 2.12). It can be noted that as extrinsic mortality approaches intrinsic mortality in magnitude (i.e. total mortality is doubled), the resulting population size declines by nearly four orders of magnitude.

These results indicate that extrinsic mortality can potentially be a major factor in determining population size. I modelled extrinsic mortality as always being *additional* to intrinsic mortality. That is, I assume that extrinsic mortality never compensates for intrinsic mortality. This may or may not be a good assumption, and this will be explored further in subsequent work. In any case, this result highlights the need to obtain estimates of this rate if we are to accurately predict the population dynamics for this species.

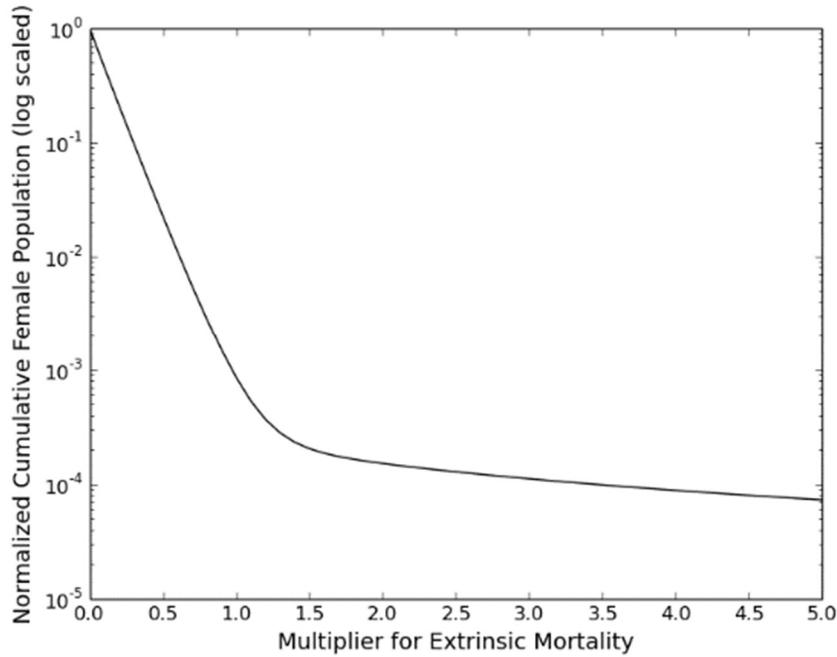


Figure 2.12 The effect of increasing extrinsic mortality in proportion to intrinsic mortality. As the extrinsic mortality increases, the cumulative population decreases.

## 2.5 Discussion

The population dynamics model presented herein has been developed using known life history traits of *D. suzukii* and its development is consistent with empirical data and known mechanistic responses to air temperature (Ryan et al. 2016, Emiljanowicz et al. 2014, Tochen et al. 2014, Kinjo et al. 2014). Temperatures that remain well below those required for optimal development produce relatively small populations due to a prolonged development period, reduced fecundity and increased intrinsic mortality. Temperatures that peak above the optimal development temperature, and that remain high, also have a negative impact on population size due to an increase in temperature-dependent mortality and reduced development rate, while ideal temperature profiles remain within a few degrees of optimal temperature for development. As expected, cumulative population numbers varied by region, often by several orders of magnitude, depending on regional temperature profiles.

I also considered the effects of diapause on population responses. Photoperiod-induced diapause has been characterized in numerous Drosophilid species (Kimura, 1990; Lumme et al., 1974; Salminen et al., 2012; Saunders et al., 1989). For example in *D. melanogaster* newly eclosed females

undergo diapause when exposed to short days (<14 h of light); photoperiod-induced diapause is then terminated rapidly by transfer to 18 °C (Saunders et al., 1989). However, there is relatively little information on diapause in *D. suzukii*. Ryan et al. (2016) found that pre-exposure to short photoperiod (10:14 L:D) had no effect on the survival responses of adult *D. suzukii* to long-duration low-temperature exposure relative to controls (15:9 L:D), however the level of ovarian maturation in that study was unknown, and as such it was not clear if short photoperiod induced physiological diapause. Jakobs et al. (2015) also found no evidence of diapause in *D. suzukii* following two weeks of exposure to simulated fall conditions, though while photoperiod may have been short enough to induce diapause (11.5:12.5, and 12:12 L:D), daytime temperatures in that study were relatively high during this exposure (9/21 °C and 5.5/19 °C). Consequently, the diapause sub model presented here is both simplified and generalized due to a lack of specific detail on *D. suzukii* diapause characteristics, but could easily be refined as data become available.

Diapause sensitivity analysis suggests that diapause termination temperature is more important for population growth than the photoperiod at which diapause is induced. This is because low fall temperatures have a strong negative impact on development and fecundity in advance of diapause induction. Empirical observations of *D. suzukii* temperature-dependent development show that fall temperatures substantially slow development. Ryan et al. (2016) found that egg-to-adult development of *D. suzukii* took 30.3 days at 15 °C, 75.1 days at 10 °C and was suspended at temperatures of 8 °C and below. Similar results for *D. Suzukii* have also been found by Tochen et al. (2014). As such, in many cases development had slowed or stopped in the current model in advance of diapause induction. However, temperature-based diapause termination could significantly impact overall population size with early diapause termination, coupled with consistently warm temperatures, representing ideal conditions for population growth. Moderate climates are the most sensitive to termination temperature, while areas with warm temperatures are less sensitive to this parameter, as the higher temperatures ensure that the population breaks diapause early and cumulative population size is maximized (Figure 2.7). These runs were performed with mathematically generated temperature profiles however, and the behaviour is likely to be different in reality, where temperatures are fluctuating. For example, in cases where diapause breaks early because of transient unseasonably high temperatures, populations may be at a disadvantage once temperatures return to lower seasonal norms (Figure 2.6). Fluctuating temperatures have also been shown to alter temperature-dependent life history traits. For example, *D. melanogaster* reared under fluctuating-temperature conditions had a higher tolerance to both heat and cold than those exposed to constant

temperatures (Overgaard et al., 2011).

Analysis of the sensitivity of the model to the initial number of individuals demonstrates that the peak population predictably scales with the initial number of females (Figure 2.6). This highlights the importance of overwintering survival to the final population size. However, the model is limited to single-year simulations due to a lack of available data on the overwintering of *D. suzukii*, though some progress has been made in this area. It is now known that *D. suzukii* are chill-susceptible and are likely to have low survival in regions that experience low winter temperatures. Jakobs et al. (2015) found that 80% of flies were killed after 1 h at -7.2 °C for males, and -7.5 °C for females. Additionally, it was found that populations in field cages in Southern Ontario were killed early in the winter by a transient cold snap (Jakobs et al., 2015). This would suggest that in severe winters, where temperatures regularly fall below zero, spring populations of *D. suzukii* are likely due to new migrations or from populations overwintering in human-made structures. As data are collected on overwintering, immigration, invasion and transportation characteristics, the model may be enhanced to produce multi-year simulations with carry-over of individuals between years. Further, for many regions humidity is also likely be an important characteristic for development fecundity and mortality (Tochen et al. 2015) but is not considered here.

The current model also illustrates the impact of fruit quality on population size. Fruit of high quality and availability can, in conjunction with environmental factors, optimize population growth while lower quality fruit (pre-ripened) and low availability (post-harvest) are likely to have a negative impact. The fruit quality sub model presented here is simplified and represents no specific fruit development model though it is general enough that it may be parameterized for specific hosts for which development data are available. Figure 2.11 illustrates the effects of host availability on population size and shows that unlimited availability of appropriate hosts can result in cumulative population numbers that are several orders of magnitude higher than situations where food is transiently available. *D. suzukii* have many known non-crop hosts that may grow in field margins, hedgerows, or areas with unmanaged woody or riparian ornamentals, shrubs or vines that are able to sustain *D. suzukii* populations (Lee et al., 2015). As such, *D. suzukii* populations in crops are likely heavily influenced by the availability of wild hosts between harvests. The impact of fruit availability on population numbers therefore strongly argues for management practices that include cultural controls such as removal of wild hosts and destruction of rotting fruit post-harvest.

Wiman et al. (2014) have produced a model to explore *D. suzukii* population dynamics with several similarities to the model presented here. Both models incorporate environmental factors,

particularly temperature, to determine survival and fecundity, both assume that food is always available to the population and both minimize early-year negative population growth using roughly analogous ‘biofix’ (Wiman et al. 2014) and diapause termination. There are however, important differences. The Wiman model is based on Leslie matrices and is thus a discrete time model, and critical population parameters are based on the empirical observations of Tochen et al (2014). In contrast, my model is continuous in time representation and is based primarily on experiments conducted by Emiljanowicz et al. (2014) and Ryan et al. (2016). While the observations of Tochen et al. and Ryan et al. are similar, experimental changes such as fruit versus artificial diet (respectively) may account for discrepancies in the resulting parameters which may in turn be reflected in the differing results between the two models. Finally, I have included a diapause sub model whereas diapause appears to be considered only during the calculation of the initial population size in the Wiman model. When the diapause and fruit sub models are included the difference in resultant population size between the two models (supplemental Figure A1.6) are amplified as the model shows a sharper decline in population in the later months due largely to the induction of diapause and the decrease in fruit quality.

Another recent model is that of Gutierrez and Ponti (cf. Asplen 2015, supplementary material). They present a physiologically-based demographic model (PBDM) parameterized with development and mortality rates based on the data of Tochen et al (2014), Dalton et. al (2011), Kinjo (2014) as well as unpublished laboratory data. The model is stage specific including egg, larval, pupal and adult stages. Although an explicit diapause phase is not included, a reproductive quiescence is initiated when food hosts are unavailable (due to low temperatures) or temperatures are considered too low to support reproduction. Unlike the model presented here, Gutierrez and Ponti include consideration for relative humidity on fecundity and overwintering of reproductively quiescent adults. Despite the differences many of the conclusions remain the same; temperature and host availability are among the principle drivers of population size and as such warmer climates are likely to experience larger and potentially more destructive infestation than cooler climates. Taken together, the differences between the three models and their predictions help shape understanding of how different abstractions add to the knowledge of population dynamics in this species.

Given that *D. suzukii* is a relatively new pest in North America, research is ongoing to establish its actual and potential impact on the soft-skinned fruit and berry industries (Goodhue et al. 2011, Bolda et al., 2010). Studies of the effectiveness on pesticides and information campaigns (Beers et al. 2011, Dreves 2011, Bruck et al. 2011, Walse 2012) are already underway. It is important to

acknowledge the cost of these measures, both financial and environmental, and that fruit-growing regions are not at equal risk for significant loss due to infestation. In some areas environmental factors may prohibit *D. suzukii* from reaching the required critical mass and might avoid costly chemical treatments. In other areas such treatments may be routinely unavoidable. However, inter-annual fluctuations in temperature and longer-term climate change may cause SWD population levels to change, which may require appropriate adjustments in management strategies. As further field and laboratory data becomes available models such as the one presented here will continue to add important insights and may be considered important tools for stakeholders looking to prevent or mitigate the effects of *D. suzukii* in high risk regions. Furthermore, as the global climate changes relative risk may also change on a regional basis. Areas of low risk due to temperate summers and colder winters may become more hospitable as temperatures rise while currently high risk areas may experience reduced *D. suzukii* presence due to an increase of temperature outside of developmental or mortality range. Regional cooling on a shorter time scale as climate shifts occur may also have an impact on the presence and impact of *D. suzukii*. These climate impacts are an important area of study when considering a long-term strategy for managing this species (Langille et al. 2017).

## 2.6 Acknowledgements

This research was supported by grants from the Ontario Ministry of Agriculture, Food and Rural Affairs (OMAFRA) and the Canadian Natural Sciences and Engineering Research Council (NSERC) to JAN. The authors would also like to thank Nik Wiman et al. for providing data and the permission to replicate it in this work, as well as Steven Zarchney for assistance with GIS.

### 3. The Impacts of Climate Change on the Abundance and Distribution of the Spotted Wing Drosophila (*Drosophila suzukii*) in The United States and Canada<sup>3</sup>

#### 3.1 Abstract

*D. suzukii* is a relatively recent and destructive pest species to the North American soft-skinned fruit industry. Understanding this species' potential to shift in abundance and range due to changing climate is an important part of an effective mitigation and management strategy. I parameterized a temperature-driven *D. suzukii* population dynamics model using temperature data derived from several Global Circulation Models (CMIP5) with a range of relative concentration pathway (RCP) predictions. Mean consensus between the models suggest that without adaptation to both higher prolonged temperatures and higher short-term temperature events *D. suzukii* population levels are likely to drop in currently higher-risk regions. The potential drop in population is evident both as time progresses and as the severity of the RCP scenario increases. Some regions, particularly in northern latitudes may experience increased populations due to milder winter and more developmentally-ideal summer conditions but many of these regions are not currently known for soft-skinned fruit production and so the effects of this population increase may not have a significant impact.

Keywords: *Drosophila suzukii*, global circulation model, CMIP5, invasive species, elevated temperatures, soft-skinned fruit industry

---

<sup>3</sup> Published in part as: Langille AB, Arteca EM, Newman JA. (2017) The impacts of climate change on the abundance and distribution of the Spotted Wing Drosophila (*Drosophila suzukii*) in the United States and Canada. PeerJ 5:e3192 <https://doi.org/10.7717/peerj.3192>

### 3.2 Introduction

Climate change is likely to alter the abundance and distribution of invasive species, via changes in: introduction pathways, effectiveness of management strategies, and altered climatic constraints (Hellman et al., 2008). Changes in key environmental factors such as temperature (Ward & Masters, 2007) will influence species' ability to survive and thrive in regions for which climate no longer acts as a natural boundary. Invasion and infestation risks may increase as shorter, warmer winters open northern regions to overwintering possibilities and to earlier and longer growth seasons. Conversely, for some species there may be a reduction in suitability as climate factors begin to exceed mortality and development tolerances. Examining the potential changes in abundance and distribution of pest species due to climate change is vital from economic and food-security perspectives (Ziska et al., 2011).

*Drosophila suzukii* is a relatively recent species-of-concern for soft-skinned fruit producers. While undocumented outside of Japan prior to the 1930s (Kanzawa, 1939), it has increased its geographic range rapidly over the past 35 years and can now be found on every continent except for Antarctica (Asplen, 2015). *D. suzukii* is particularly concerning because unlike other “vinegar” flies its eggs are deposited preferentially into ripe (or ripening) fruit through the female’s serrated ovipositor (Atallah et al., 2014). Host fruit then spoils due to larval feeding or secondary infection at the puncture site (Goodhue et al., 2011). It is estimated that production losses due to *D. suzukii* infestation could reach 20% of crop yield or above \$5 million (USD) from preferred hosts including strawberries, blueberries, cherries, blackberries and raspberries.

As with most insects, temperature is among the critical factors influencing *D. suzukii* population size and consequently infestation potential. Several laboratory studies have highlighted the role of temperature in development and mortality and as such baseline physiological data is available. For example, Ryan et al. (2016) found that mortality thresholds occurred at 5°C (lower) and 35°C (upper), that no adult eclosion occurred at temperatures less than 8.1°C or above 30.9°C and that optimal temperatures for development and reproductive output occurred at 28.2°C and 22.9°C respectively. Similarly, Tochen et al. (2014) found the optimal rate of population growth at 22°C and development rates occurred at minimal, optimal and maximal temperatures of 13.4°C, 21.0°C, and 29.3°C respectively.

Langille et al. (2016) have developed a mechanistic model of *D. suzukii* population dynamics (Figure 3.1, see also Gutierrez, 2016), including sub model considerations for reproductive diapause and influence of fruit quality or viability. The principle environmental driver for the model is temperature and key biological mechanisms such as mortality, fecundity and development rates (for each life stage) have been parameterized wherever possible using the laboratory data of Ryan et al. (2016) and other literature-based estimates (Emiljanowicz et al. 2014). The model does not currently account for overwintering so all simulations run for a single year and typically begin with an initial number of adult females introduced on the date when diapause would be terminated due to adequately warm temperatures. Parameterizing the model with observed temperatures from 1993–2013, Langille et al. demonstrated potential relative population levels for twelve berry producing regions across Canada and the United States. The goal of that work was to demonstrate that, assuming temperature is a key determinant of population size, not all fruit-producing regions are at equal risk for infestation potential or severity, despite the almost ubiquitous presence of *D. suzukii* across North America. My goal in the present work is to explore how infestation-potential may be affected due to projected climatic change.

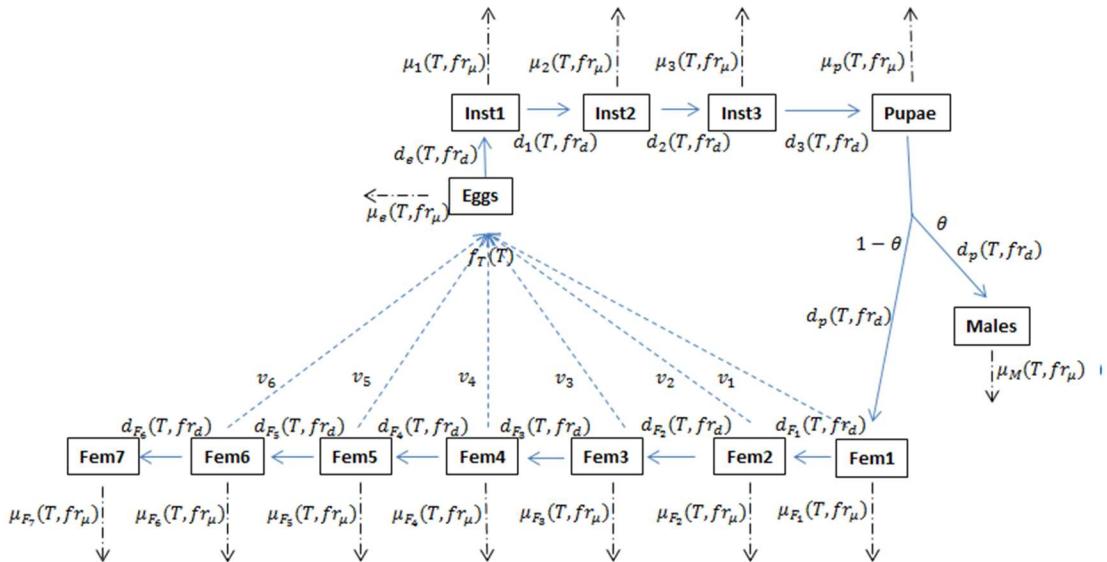


Figure 3.1 Schematic of the *D. suzukii* population dynamics model presented in Langille et al. (2016). This schematic highlights population stage structure and relationships between state variable equations including principal input and output parameters. Fecundity ( $f$ ), development ( $d$ ) and mortality ( $\mu$ ) processes are temperature ( $T$ ) dependent while development and mortality also rely on the fruit quality sub model ( $fr$ ). In addition to eggs, pupa and adult males, the model contains three juvenile instar stages and seven adult female stages in order to account for variation in egg viability ( $v$ ).  $\theta$  represents the ratio of males to females and is typically set at 0.5. Solid arrows indicate developmental stage transition, dashed arrows indicate fecundity and dot-dashed arrows indicate mortality.

### 3.3 Materials and Methods

Projecting future climate on a global scale is an exceptionally complex task given the number of processes and interactions that occur in the biotic and abiotic environment. Despite the challenges, Global Circulation Models (GCMs) continue to be refined and offer projected climate based on various modelled assumptions about those processes and interactions. The most recent suite of models follow the Coupled Model Inter Comparison Project phase 5 (CMIP5) framework and differ from their predecessors, in part, in their use of Representative Concentration Pathways (RCPs). Each of the RCPs represents a potential future trajectory of the principle forcing agents (measured in W/m<sup>2</sup>) driving climatic change. In simplified terms RCP2.6 represents a “peak-and-decline” scenario where radiative forces peak by mid century and decline by 2100. Both RCP4.5 and RCP6.0 are scenarios where stabilization of radiative forces occur by 2100, while in RCP8.5, these forces continue to increase beyond 2100 (van Vuuren, 2011). A subset of possible temperature projections was downloaded from the archives of the Downscaled CMIP3 and CMIP5 Climate and Hydrology Projections (<ftp://gdo-dcp.ucllnl.org/pub/dcp/archive/cmip5/bcca>). Specific climate models were chosen where projections were available for each of the four Representative Concentration Pathways (RCPs) 2.6, 4.5, 6.0, 8.5. Selected climate models included the BCC-CSM1.1, CCSM4, GFDL-EMS2G and ISPL-CM5A-LR (Table 3.1), all of which conform to the Coupled Model Inter Comparison Project Phase 5 (CMIP5) framework.

Modelling Center or Group	Institute ID	Model Name
Beijing Climate Center, China Meteorological Administration	BCC	BCC-CSM1.1
National Center for Atmospheric Research	NCAR	CCSM4
NOAA Geophysical Fluid Dynamics Laboratory	NOAA GFDL	GFDL-EMS2G
Institut Pierre-Simon Laplace	ISPL	ISPL-CM5A-LR

Table 3.1 Official contributor name, institute ID and model name for the GCMs used herein.

The temperatures obtained were derived from projections downscaled via Bias Corrected Constructed Analogs (BCCAv2, Maurer et al., 2007; Brekke et al., 2013) and consist of daily minima and maxima at 1 degree (latitude/longitude) spatial resolution. I computed daily mean as (minimum + maximum) / 2. Daily data were then averaged over thirty-year time spans such that three time periods 2020s (2010-2039), 2050s (2040-2069) and 2080s (2070-2099) were produced. Each simulation began with 10 fecund females introduced on the date for which temperature would break

reproductive diapause ( $18^{\circ}\text{C}$ ). The fruit development index which controls the rate at which fruit reaches maturity was set to 4.0 and fruit was available at maximum quality for 50 days before beginning to decline. Multiple harvests per year were not included in these simulations. Reproductive diapause is initiated when the simulated grid cell reached 10 daylight hours per day (latitude-dependent). For a complete list of model parameters please consult Langille et al. (2016, supplemental material).

Description	Value
Initial population of fecund females (individuals)	10
Diapause termination temperature ( $^{\circ}\text{C}$ )	18
Diapause induction date based on daylight hours	10
Fruit development index ( $\omega$ )	4.0
Fruit harvest lag ( $h_{Fr}$ , days)	50

Table 3.2 Key base model parameters for simulations, *cf.* Langille et al. (2016, supplemental) for complete parameter value listing.

As mentioned in section 2.3.7, the base population dynamics model was implemented using C++, specifically GNU gcc version 4.8.4 for Ubuntu Linux (14.04.03 LTS). All differential equations (eqn 15 through eqn 19) were solved using Euler's method of numerical integration. Source code files including headers are freely available at <https://github.com/alangillGuelph/dsPopSim>.

### 3.4 Results

In Figure 3.2 I illustrate the base model properties under a subset of temperature scenarios from the CCSM4 model for Burlington, New Jersey. The temperatures for three different RCPs and time frame combinations (Fig 2.a) show a relatively low (RCP2.6-2020), mid (RCP6.0-2050) and high (RCP8.5-2080) range of potential conditions. The difference in the temperature profiles result in varying fruit readiness date (Fig 2.b) and population levels (Fig 2.c). In particular, RCP8.5-2080 produces the earliest optimal fruit in part due to warmer temperatures early in the year while fruit ripens more quickly in RCP2.6-2020 than in RCP6.0-2050, due to slightly elevated temperatures in the March through May range. It is worth noting that the RCP6.0-2050 temperatures are generally higher than

those of RCP2.6-2020 in the later part of the year (June and onward) but this occurs after fruit has matured and been harvested. Despite some variation between the curves, there are only 14 days difference between the earliest and latest readiness dates.

In terms of population (Figure 3.2c), RCP8.5-2080 initially grows at the steepest rate but peaks relatively early in the year and begins to decline due to temperatures (Figure 3.2a) that are both too high to maximize development and contribute to mortality. RCP6.0-2050 growth is slowest due to the lower early-year temperatures which also causes slower fruit quality increases (Figure 3.2b). In comparison, RCP2.6-2020 provides maximum growth potential with early warmth and lower mid-to-late year temperatures that continue to foster growth without severely impacting mortality. Considering both reproductive output and development, modelled population growth occurs optimally at temperatures around 24°C. Quantitatively, temperature profiles for RCP2.6-2020, RCP6.0-2050 and RCP8.5-2080 produce 74, 62 and 47 days respectively where temperatures are  $\pm 2^{\circ}\text{C}$  from model optimal.

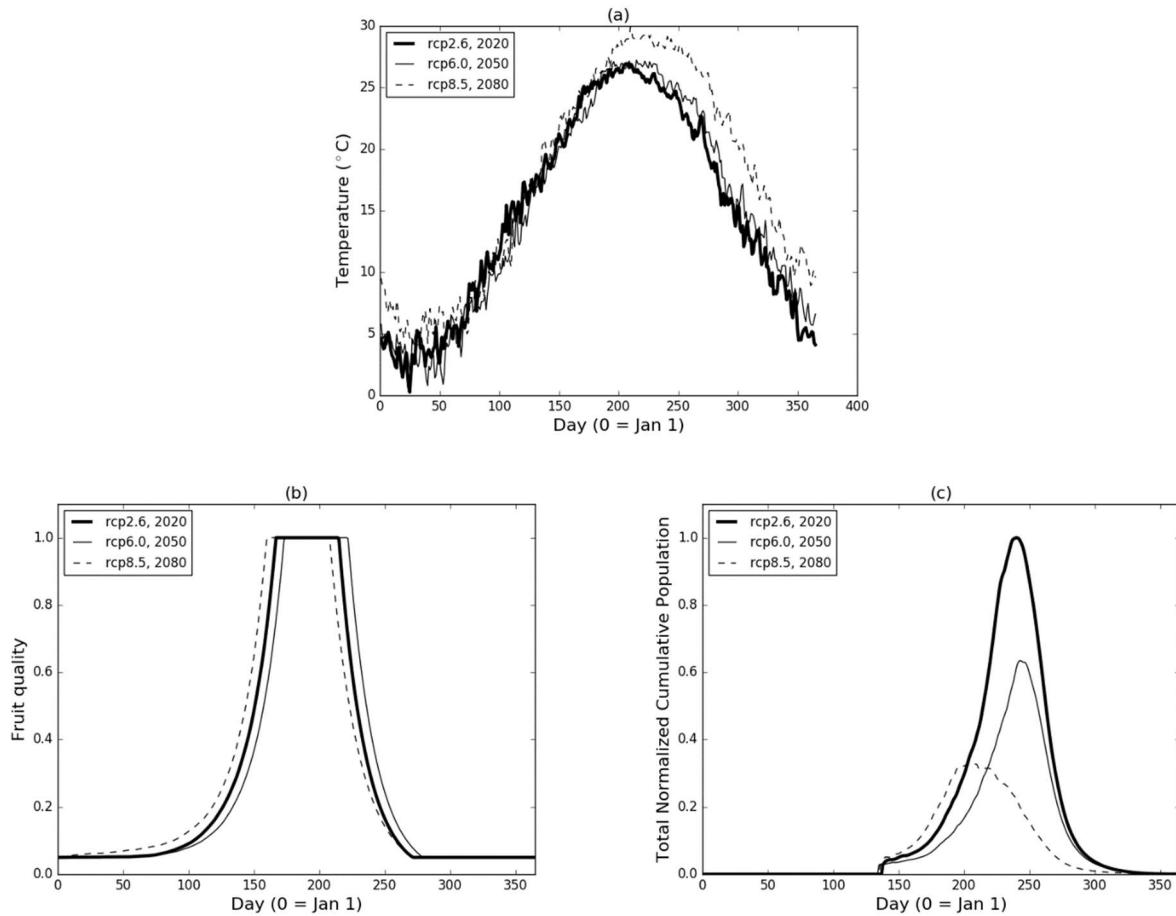


Figure 3.2 a) Sample temperature profiles for Burlington, New Jersey from the CCSM4 model with different RCP-date combinations. Model-RCP combinations were selected to illustrate a range of output possibilities. b) and c) illustrate fruit quality sub model response and model population based on temperature profiles in a).

As it cannot be known *a priori* which of the climate models will produce the most accurate temperature projections, I combined the results to form ‘consensus maps’ (Fig 3). Simulated population totals (all life stages) were normalized across the entire data set, and means were generated per corresponding RCP and timeframe (i.e. these are averages of the results of the *D. suzukii* model, run for the different climate models). Based on these consensus maps, the 2020s have the potential for the most extensive high-population distribution across southern Canada and much of the continental United States. Northern Canada and regions occupied by the Rocky Mountains have relatively low populations due to lower temperatures that discourage development and encourage mortality. Similarly, the southern extent of the United States has relatively low populations

due to excessively high temperatures. There is a noticeable shift in population distribution moving to the 2050s as the central US becomes warmer and populations decline. Some northern central regions in Canada increase in population as temperatures become more favourable. These trends continue into the 2080s as the central and northern US become too warm, central Canada becomes warmer and the most favourable conditions in North America are located primarily to the north-eastern and western US as well as a few regions in southern Canada. While there is little qualitative difference between the model output in the RCP scenarios for the 2020 temperature projections, by the 2080s there is an evident decline in relative population levels in RCP 8.5.

As with the climate model, it is difficult to predict which RCP scenario will produce the most likely greenhouse gas emissions, so I further aggregate the model results by folding the model-RCP combinations into per-time frame consensus maps. Similar to Figure 3.3, in Figure 3.4 (upper) a general reduction in population sizes can be seen with the exception of moderate increases in the far northern regions as the results move into later time frames. While these mean populations give an overall impression of possible trends it is important to acknowledge the difference in values produced by the various model-RCP combinations.

Figure 3.4 (lower) illustrates these differences by showing the coefficient of variation between the normalized population values. For the 2020 model-RCP combinations relatively low values occur across the map. The exceptions are larger values to the extreme south in Florida, California and Texas. The 2050s show a larger range and an increase in high-variation locations, particularly in the extreme south and north. These trends continue into the 2080s, where larger variations occur over larger areas to the north south and west. Despite these large areas of variation, there remain areas of moderate and even low variation particularly in the interior.

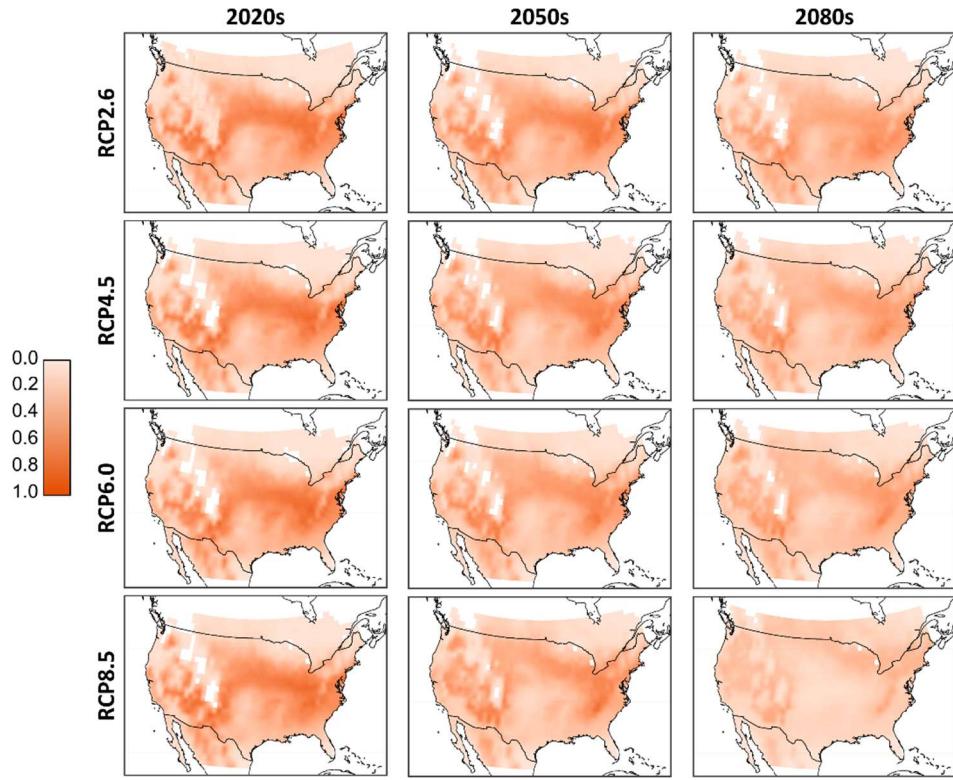


Figure 3.3 Aggregated consensus maps of the modelled populations. Data are the computed mean normalized total population (all life stages) from all four climate models. Normalization is based on the entire pre-mean data set. White cells located in the interior parts of the maps reflect missing values in the downscaled climate model outputs and a 0.25 degree bi-linear interpolation was applied to all data cells.

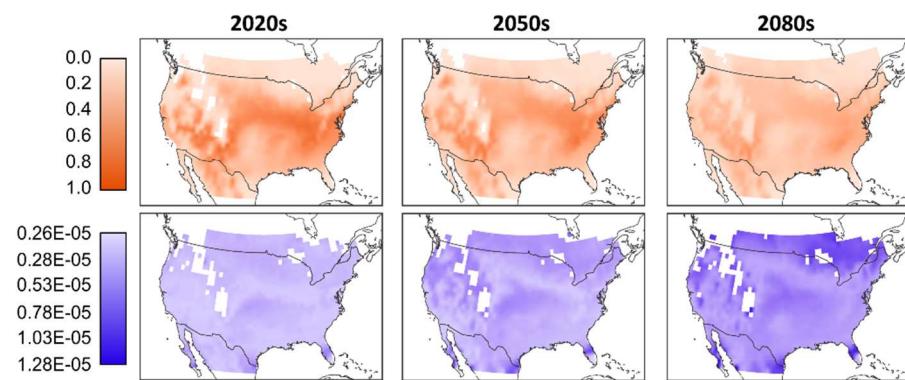


Figure 3.4 (upper) Mean normalized population distribution across North America per-timeframe. Normalization occurs across the entire data set and mean is across all models and RCPs. (lower) Corresponding per-timeframe coefficient of variation across all models and RCPs. Coefficient of variation is based on the normalized population data. White cells located in the interior parts of the maps reflect missing values in the downscaled climate model outputs and a 0.25 degree bi-linear interpolation was applied to all data cells.

Figure 3.5 provides an alternate illustration of the overall uncertainty (i.e., larger coefficients of variation) among the model-RCP combinations. It shows both a negative correlation between coefficient of variation and mean population size and an increase in variation as the timeframe increases. This suggests that the largest per-cell uncertainty among model-RCP temperature estimates are in regions that would typically produce smaller populations (too warm or too cool) and that these estimates continue to diverge as the timeframe increases.

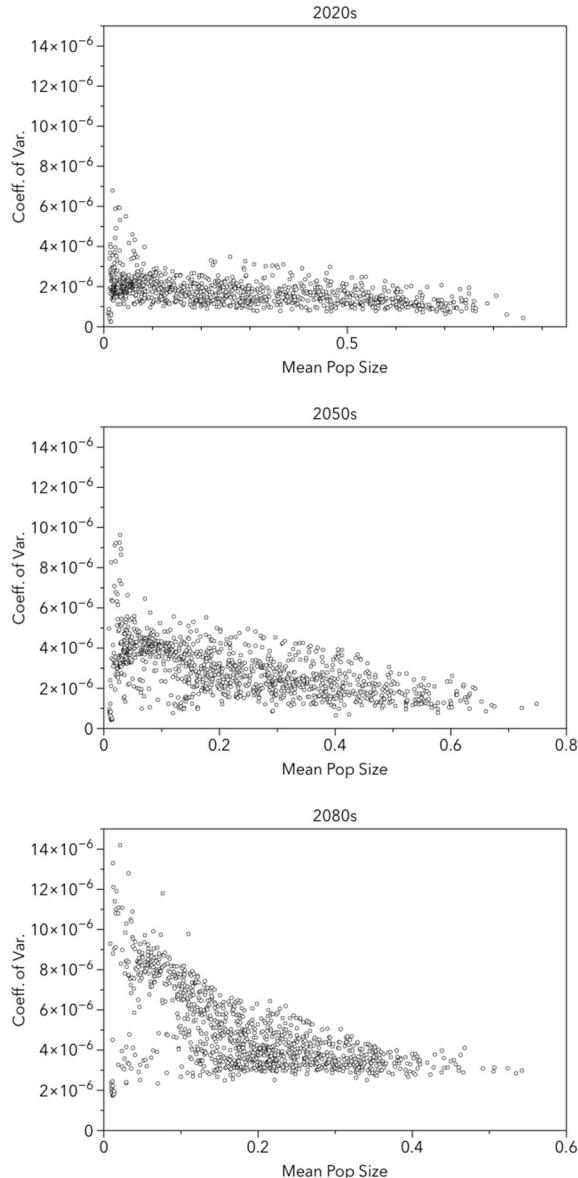


Figure 3.5 Grid cell coefficient of variation in population sizes versus the grid cell mean population size. The uncertainty in the population sizes that result from the differences between the climate model–RCP combinations tends to be larger as the mean of those population sizes gets smaller. Not surprisingly, the uncertainty is also larger later in the century than earlier.

In Langille et al. (2016) I presented potential modeled populations for several important berry producing regions in the United States and Canada. Figure 3.6 illustrates how the GCM model and RCP scenarios may affect population simulations in comparison with historic temperatures. For each location mean daily temperatures from 1993 through 2013 were calculated. More than half of the locations including Bladen North Carolina, Burlington New Jersey, Wayne Mississippi, Hillsborough Florida, Santa Barbara California and Clark Washington, have historic populations higher than all projected averages. This suggests that these areas are currently in or near the time frame for maximum population growth potential and that as temperatures increase over time populations should decrease. For the remaining locations including Allegan Michigan, Norfolk Ontario, Chicoutimi Quebec and Fraser Valley British Columbia the historic population is consistently lower than all of the projected averages suggesting that growth potential has yet to be maximized in those locations. It is also worth noting that most of the populations are projected to decrease by the end of the present century, including Allegan and Norfolk. The exception to this decline are the two northern most regions (Fraser Valley and Chicoutimi) which may continue to increase though their population levels remain relatively small.

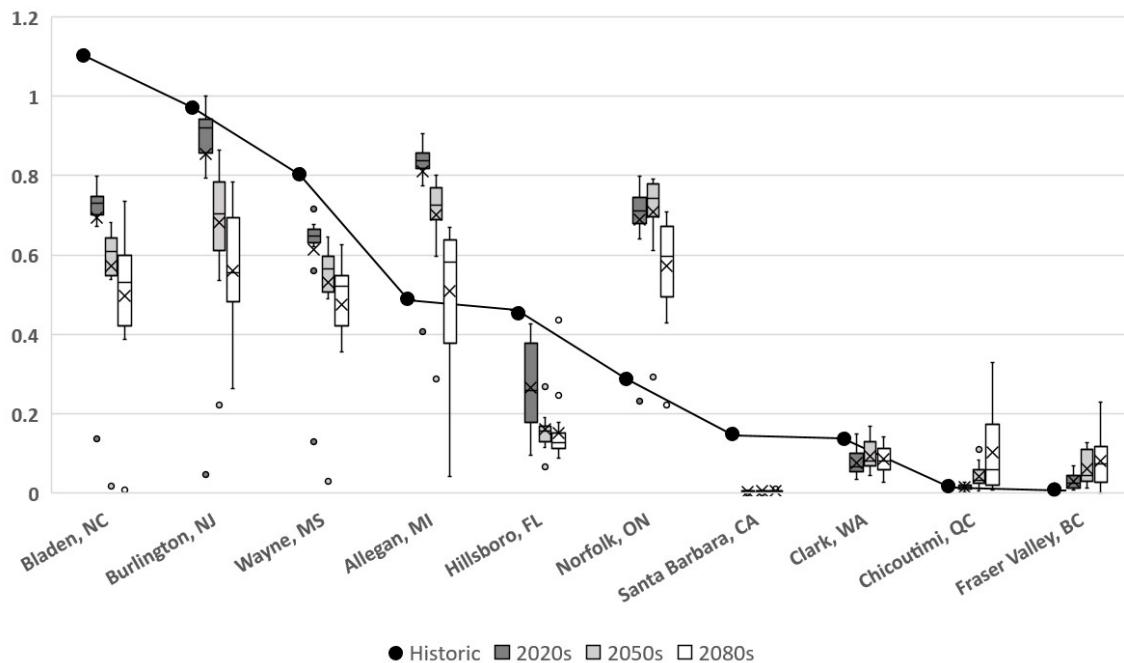


Figure 3.6 Normalized populations for historic temperatures and GCM projections. GCM projections are averaged over all climate model-RCP combinations where  $\times$  denotes mean population, whiskers denote maximum and minimum values, interior bar marks quartiles and small circles denote outliers. For averaged historic temperatures, large circles represent population normalized to GCM results.

### 3.5 Discussion

Understanding how pests such as *D. suzukii* may shift in range and density over time is an important part of an effective mitigation and management strategy (Asplen et al., 2015, Dreves, 2011). By combining various climate models and emissions scenarios I provide a mean outlook as to how increasing temperatures may affect infestation potential over the coming decades. The results of the model simulations offer a somewhat mixed forecast for *D. suzukii* infestations. They suggest that some regions in North America may already be at or near the peak of their population growth potential. Distributions and infestation hot-spots may shift over the next two or three decades but as toward the end of the century, depending on the degree of climatic change, there may be a marked decline in population growth potential from this destructive pest as temperatures become too warm to sustain development and mortality increases. Beyond sustained higher temperatures, an increase in short-term heat events in susceptible regions may adversely affect crop output (Teixeira et al., 2011) and subsequently lower pest population potential. In contrast, some northern regions may become more favourable for *D. suzukii* population growth and although many of these regions are not currently known for commercial soft-skinned fruit production currently this too may change over time (Olesen and Bindi, 2002).

It is important to address the simplifications of the model used in this work. At this time, there is no consideration for overwintering and as such the model is limited to single year simulations. Although empirical work continues to highlight the biological mechanisms and consequences of overwintering (Wallingford et al., 2016, Wallingford and Loeb, 2016, Shearer et al., 2016, Stephens et al., 2015) it has not yet been included in the model (see discussion in Langille et al. 2016). As a result, these simulations do not account for inter-annual population differences due to winter survival rates and subsequently presume that all regions begin with the same starting population size each year, which does not specifically reflect migrations (immigration or emigration) due to environmental or anthropogenic factors. Also, considerations for other mortality and development influencing factors such as humidity are not included but are likely to impact population growth (Tochen et al., 2016) while others such as predation and (parasitism) have been greatly simplified to reduce complexity.

Perhaps the most important limitation of the current work is the lack of accounting for evolutionary adaptation that may well occur over the course of the century. While this simplification may be reasonable when considering both the 2020s and historical temperatures, it may be more

reasonable to assume that some adaptive changes are likely to occur before the 2050s and 2080s. Both Umina et al. (2005) and Gilchrist et al. (2004) have demonstrated that rapid climate-driven physiological and adaptive change is possible in other species (*D. melanogaster* and *D. subobscura* respectively) although Kimura (2004) found no difference in thermal tolerance between cool and warm latitudinal strains of *D. suzukii* in its native region of Japan. Previous work has also shown that Drosophila species (such as *D. melanogaster*) can be cold and heat-hardened if short term temperature changes are non-lethal and that over longer term exposures acclimation can occur (Wallingford et al., 2016, Hoffman et al., 2003). Heat hardening may also increase desiccation resistance (Hoffman 1990). What remains unclear at this time are the exact thresholds for which evolutionary adaptations will protect species such as *D. suzukii* from temperature and humidity extremes and the developmental and reproductive trade-offs that may occur from such adaptations.

Although it is standard practice to average temperatures over larger time spans when using climate model data, the averaged data dampens short-term drops or spikes in temperature that might adversely affect population sizes. As such, the simulation results presented herein offer insight into the average potential for growth but do not highlight the extent of inter-annual variation that may occur. To illustrate the effect of using mean temperature data with this model, Figure 3.7 (a and b) shows a set of population curves for the historic temperature profiles (annual temperatures from 1993-2013). I also show the average of the population curves and the population when the simulation is run using the average temperatures over the 20-year time period. In some cases, such as with Hillsborough Florida, averaging the temperatures creates a near-optimal profile by greatly reducing the number of days where temperatures are either too high or too low to support development. The resulting population is larger than any of the individual observed temperature profiles would otherwise suggest. Conversely averaging the temperatures at Fraser Valley, British Columbia, results in a profile where most days are well below optimal growth temperatures. The individual historic temperature profiles, while more variable than the mean, offer the population more opportunity for growth. The resulting population from the averaged temperature profile is significantly smaller than most of the individual profiles.

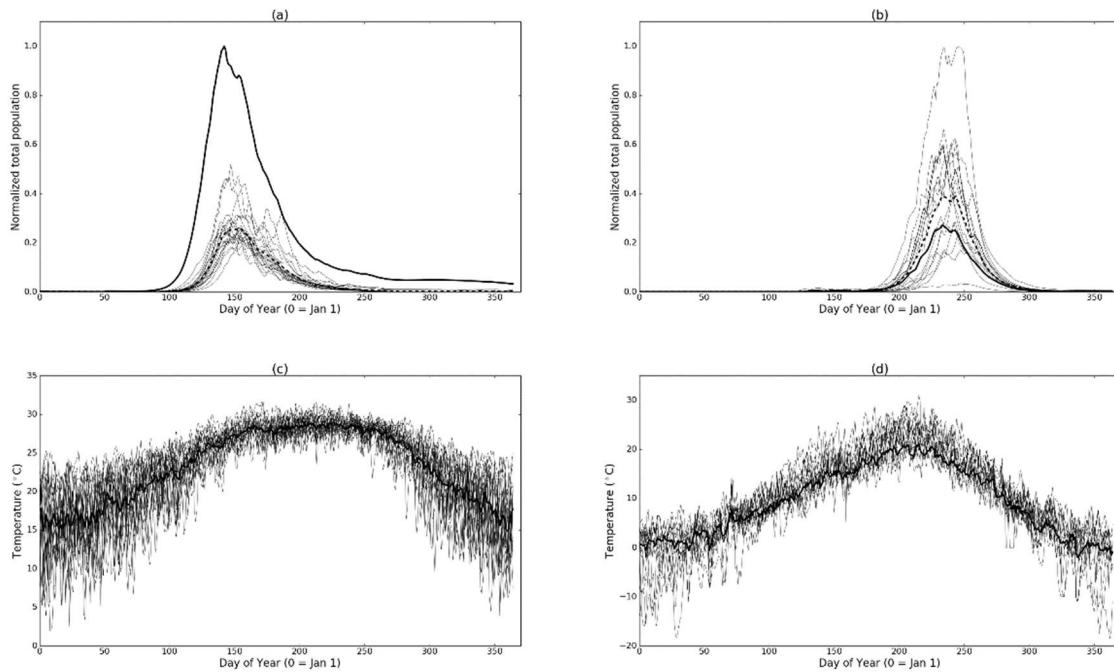


Figure 3.7 a) and b) Normalized populations over 20 years of historic data (1993 through 2013) for Hillsborough, Florida and Fraser Valley, British Columbia respectively. The dashed line indicates the mean annual population while the bold-solid line indicates the population for the mean temperatures across all years. c) and d) illustrate the daily temperature data for all years and the mean daily temperatures across all years for Hillsborough, Florida and Fraser Valley, British Columbia respectively.

Given the inherent uncertainties in both the base population dynamics model and the GCM-derived temperature data, the work presented here is not intended to be quantitative predictions. Rather it is hoped that these results, when considered with comparable modelling efforts on *D. suzukii* or similar pest species, will help to guide and refine empirical and theoretical work on the mechanisms that drive population levels and highlight the range of possible outcomes as we enter a critical time of rapid climatic change. As models continue to highlight the level of understanding of population-driving processes, empirical studies will continue to provide new and increasingly accurate quantitative inputs. Similarly, as GCMs continue to develop and understand of complex global environmental processes matures, opportunities for explorative and potentially predictive modelling of invasive species will also improve.

### 3.6 Acknowledgements

The authors acknowledge the World Climate Research Programme's Working Group on Coupled Modelling, which is responsible for CMIP, and thank the climate modelling groups (listed in Table 3.2) for producing and making available their model output. For CMIP the U.S. Department of Energy's Program for Climate Model Diagnosis and Intercomparison provides coordinating support and led development of the software infrastructure in partnership with the Global Organization for Earth System Science Portals. This research was supported by grants to J.A.N., from the Ontario Ministry of Food, Agriculture and Rural Affairs, and from the Canadian Natural Science and Engineering Research Council. The authors would also like to thank Marie Puddister and Adam Bonnycastle for their assistance with ArcGIS.

## 4. Model Design and Use

### 4.1 Abstract

Effective communication of model design, implementation and execution details are of genuine importance. In this chapter, I present these details such that the *D. suzukii* population dynamics model may be verified, used or extended by third parties. I present this information following a modified version of the standardized Overview, Design Concepts and Details protocol proposed by Grimm et al. (2006, 2010). I also include a discussion of parsimony versus mechanistic-richness as it relates to the complexity of the presented model. I conclude with a “user manual” and a brief overview of a web-based version of the model.

### 4.2 Introduction

In order to be scientifically relevant models must be presented such that any ambiguities in implementation or interpretation are minimized. Given their lengthy history as scientific tools and often succinct nature, mathematical models are relatively simple to convey. With the proper equations, parameter values and an understanding of how to interpret the results, researchers are able to reproduce the modelled system in order to validate or perhaps refute the original findings. In principle, the same should apply to computational models but too often they appear to fall far short. Unlike the universality of mathematics, computer-based models may be written in dozens of programming languages, each with specific syntax and technically-confounding details. Additionally, it is common that as model complexity increases so does the amount of computer code that must be written. Many models are simply too long to include directly in publication and lengthy computer code may not be well suited even for appendices. These issues have historically lead to poor presentation of computer-based models but there are some emerging standards that offer some guidance.

## 4.3 Overview, Design Concepts and Details (ODD)

One such standard is the Overview, Design Concepts and Details (ODD) protocol (Grimm et al. in 2006). Originally intended to address specific issues in describing individual or agent-based models, by the time the protocol was reviewed and revised (Grimm et al. 2010) more than 50 publications had implemented it. Although Grimm et al. (2010) state that the protocol “may be” appropriate for a general class of moderately to high complexity models it is unclear from their survey how many, if any, of the 50+ publications were not individual-based. It is clear that the ODD is specifically tailored to individual-based models but it also provides some insight into possible gaps in the presentation of the base *D. suzukii* population model (Chapter 2) which I endeavour to close here. The points below follow the ODD generally though some of the details have been altered to better reflect the specifics of the population model. Finally, due to the level of detail already provided in previous chapters I do not attempt to create a complete standalone description document as the protocol suggests but rather it is meant to be a supplemental source of detail for those wishing to replicate or extend the model.

### 4.3.1 Purpose

The *D. suzukii* model was developed to explore “what-if” scenarios related to population dynamics and environmental conditions. Specifically, it illustrates potential for population growth under varying temperature profiles. The principle assumption for the design of the model is that temperature is a driving factor in *D. suzukii* fecundity, mortality and development rates. The goal of the model is to identify regions that display the most favourable conditions for population growth as they are likely to be most susceptible to infestation and crop damage.

### 4.3.2 State Variables, Scale and Entities

The base model consists of 13 state variables used to track the number of individuals in each life or developmental stage of *D. suzukii* (egg, three juvenile instars, pupa, seven adult female stages

representing different rates of fecundity and an adult male stage). Other base model state variables include the time step representing the day of the year and current temperature in degrees C. When in use the diapause sub model includes the current number of daylight hours and the fruit quality sub model includes the current fruit quality index (dimensionless), and the number of days remaining in the fruit harvest lag. The model is not spatially explicit. Further detail on the state variables and parameters are provided in Chapter 2 and Appendix 1.

Model entities is an important concept in individual-based modelling (and are consequently included the ODD protocol) as it often describes the lowest level of object-oriented programming construct used during implementation. As the *D. suzukii* model is not individual-based its entities are not individual eggs, juveniles or adults. Rather the population information is aggregated into the state variables described previously. However, by approaching these entities from a programming perspective they can still be discussed which may offer some insight into the modelling process. A simplified UML-diagram of the principle model objects is shown in Figure 4.1.

The *D. suzukii* model is broken down into several different objects in order to mitigate programming complexity. In terms of model functionality, the core object is SWDPopulation. This object is responsible for development of the population through all life stages and includes both the diapause and fruit quality sub models. Given the number of parameters required by the population object, a separate object was developed to simplify the parsing of data from file and passing the parameter values and temperature data to the population object. The SWDCellSingle object manages the interaction between the population and parameter objects and is responsible for producing the simulation output for the user. The SWDSimulatorSingle object is a “runner” object that provides a simplified interface for the user to provide simple arguments that are passed to the lower level objects. SWDCellMulti uses many instances of SWDCellSingle in order to run several population simulations simultaneously. This object makes used of C++ threading to parallelize these tasks and to increase the rate of execution. Similar to SWDSimulatorSingle, SWDSimulatorMulti is a simplified user interface used to control the multi-cell object.

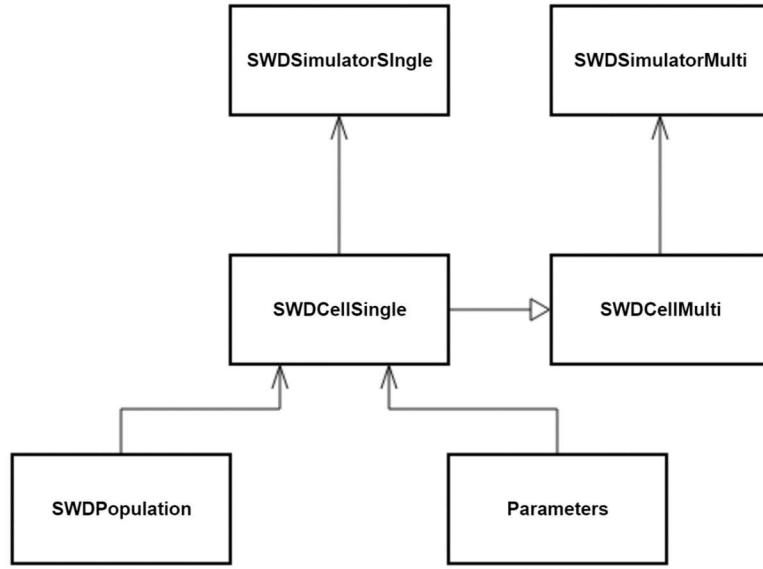


Figure 4.1 Simplified UML diagram of model object relationships. All relationships are simple association except SWDCellMulti that extends SWDCellSingle.

### 4.3.3 Process Overview and Scheduling

As the population model is not individual-based there are no complex scheduling or parallel event considerations that need to be resolved. Euler's method of numerical integration is used to subdivide the otherwise continuous nature of the model into discrete time steps. At each time step environmental variables such as temperature, and in the case of the diapause sub model, number of daylight hours, are read or computed. This information is then passed to the life stages in order and population levels for each stage are updated accordingly. Stages must be processed in order as each receives from the stage below it the population that has successfully transitioned out of that lower stage. The adult female stages provide new population back into the initial egg stage. A simplified diagram of the per-time step data flow is shown in Figure 2. It should be noted that temperature is required by the fruit quality sub model in order to calculate the fruit quality index. Further information on the parameters and values required by each stage can be found in Chapter 2 and Appendix 1.

The multi-cell version of the model has increased complexity in terms of scheduling due to the simultaneous (threaded) execution of several single-cell population simulations. However, since

there is no transfer of population or data between cells each operates independently per Figure 4.2. The order in which the individual cells in a multi-cell simulation begin or complete execution has no impact on the output data.

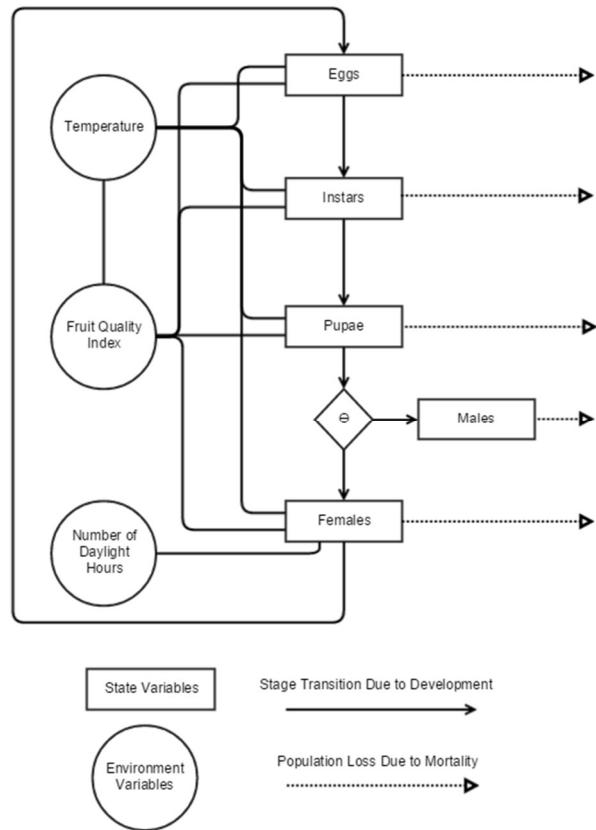


Figure 4.2 Simplified information flow diagram of the base model including diapause and fruit quality sub models.  $\Theta$  determines the ratio of females to males.

#### 4.3.4 Design Concepts

The updated ODD specification (Grimm et al. 2010) highlights several important design considerations that should be communicated with respect to model development. Many of these concepts including emergence, adaptation, learning, sensing and collectives are particularly suited for individual-based models and do not apply here. The remaining design concepts are summarized below.

*Stochasticity:* There are no stochastic elements in the presented population model. Rather, the model is fully deterministic such that for each distinct set of input parameters a corresponding and consistent set of output is produced.

*Observation:* Model output includes the population (number of individuals) for each life stage at each time step (day) over the duration of a simulation (typically, 365 days) as well as the fruit quality index per time step. In addition, summary data is also provided including total population (cumulative sum) for each life stage over the simulation time period, peak population date, peak fruit quality date and diapause termination date.

The “raw” data produced from the model reflect the exact computed population levels. As the model is not spatially explicit and population thresholds required for damaging infestations to occur remain unknown or unpublished, results presented in Chapters 2 and 3 are normalized. This normalization provides a relative measure of population size between cells, or from a modelling perspective, between input parameter sets including starting population size, temperature profile and more.

*Parsimony:* Though not included in the official ODD protocol parsimony is an important design consideration particularly for complex models, mathematical or computational. The genesis of this principle is often referred to as Occam’s razor which states that the simplest hypothesis that does not contradict the data is the preferred hypothesis. Similarly, Warren and Seifert (2011) explain “parsimony dictates that our models should be no more complicated than need be to explain our observations” (see also Bossel 2007). Simpler models, by definition, have fewer parameters, equations or sub models which often makes them easier to present which, in turn, promotes validation and reproducibility. From a computational perspective, simpler models may require fewer resources such as computational time or power which may allow for broader exploration of parameter space or simplified sensitivity analysis. With the accessibility of modern computing power, this particular justification for model simplification seems somewhat outdated. In general, these practical benefits suggest that parsimony is a desirable objective during model development and a valid criterion when selecting between competing models (March and Hau 1996, Preacher 2006). However, there is a substantial body of literature indicating that the limitation of parameters or sub models must be considered in conjunction with both goodness-of-fit to data and model objectives. In particular, models that maximize fit may be unnecessarily complex while the maximization of

parsimony may sacrifice fit (Johnson and Omland 2004).

March and Hau (1996) evaluated the use of parsimony indices to guide selection amongst competing structured equation models. They find that even when the difference of fit between models is quantitatively small there may subjective cases where the non-parsimonious model proves more effective in representing the modeled system. They stress the importance of researcher intuition and experience and suggest that other criteria beyond simple parsimony may be more effective in model evaluation. Houlahan et al. (2016) question the traditionally strong emphasis on model parsimony by suggesting that its value is not inherent in simplicity itself but rather in the practicality of obtaining and managing higher dimensional parameter estimates. Reichert and Omlin (1997) claim that the elimination of processes and parameters can artificially reduce uncertainty in the model and consequently underestimate the range of results in predictive models (see also Jakeman et al. 2006). Their stance is that this uncertainty (specifically in ecological models) should be part of the model analysis and can only be accurately gauged when knowledge of existing processes is included even if this means the development and use of “overparameterized” models.

DeAngelis and Mooij (Lauenroth et al., 2003) specifically advocate for increased development of mechanistically-rich models (see also Topping et al. 2010) stating that simplicity should not be a design or development goal at the expense of “explanatory power”; ecological systems are often complex and simple models often do not adequately describe the operational processes. They describe and refute several traditional criticisms of mechanistically complex models. The first criticism is the larger volume of data typically required for both process representation and parameterization. They contend that data requirements can typically be satisfied if models are tailored to available empirical data, data needs are limited to simplified representations of processes and multiple data sources are used. For economically important species, like agricultural pests, the data are more likely to be available, than for natural systems more typically studied by ecologists. The second criticism often associated with mechanistically rich models is that the inherent complexity exacerbates errors and uncertainty. DeAngelis and Mooij argue that the access to intermediate processes allows modellers to mitigate errors with greater ease and specificity than simpler process-aggregated models. The final criticism addressed is that complex models lack the descriptive transparency of their simple counterpart to which DeAngelis and Mooij offer two perspectives; while a model might be complex overall the individual processes may be relatively simple to convey and the biological realism of the individual processes promotes “conceptual clarity”.

As highlighted by DeAngelis and Mooij a model’s explanatory power may be reduced through

aggressively parsimonious design and as such parsimony may not be an effective criterion for evaluating between competing models when explanatory power is desirable. In many cases parsimony remains a valid criterion for model evaluation, in other cases it may be counter-productive and should be considered in context of model objective (see Chapter 1).

All models are abstractions of the system that they are meant to represent. The modeller has to make decisions about which processes are critical, and must be represented mechanistically, which processes are important, but can be simplified and represented phenomenologically, and which processes are probably unimportant, and can be left out entirely. There aren't right and wrong decisions about which abstraction is the 'best' abstraction. 'Best' is intimately linked with purpose, and the reader's purpose for reading about the model may differ substantially from the modeller's purpose in constructing the model. Parsimony can be understood as a particular take on what is the 'best' level of abstraction, but it is not the only position, nor necessarily the right position. In this context, it is probably worth recalling the purposes behind the construction of this particular model.

From Chapter 2:

"We have two principle objectives in constructing this model: (i) to gain a better understanding of the temporal and spatial patterns of *D. suzukii* population growth, and hence risk, that result from particular assumptions about crucial mechanisms; and (ii) to create a useful tool for answering interesting "what if" questions regarding *D. suzukii* invasion, climate and climatic change. A mechanistic model can help us to meet both of these objectives. Mechanisms are a critical part of biological understanding. "What if" questions usually involve changes in the environment, autecology, or synecology of the organism. These questions require changes to the model system that are not readily accommodated in non-mechanistic approaches.

By mechanistic we mean that the model is based upon what we think are the most important mechanistic processes that determine population dynamics. Such models are particularly useful because they allow any model prediction to be traced back to the process(es) that most influences it. Since mechanistic models often have many parameters, all of which are tunable, our goal is only to achieve satisfactory quantitative behavior. Our goal is not forecasting, for which statistical models are often better suited (see e.g., Thornley and France, 2007; for further discussion)."

With these comments in mind, it is worth examining some of the modelling decisions that

were made with respect to the population model outlined in Chapter 2. The model consists of 13 differential equations to express the specific life stages of *D. suzukii* and several supporting equations to describe fecundity, development and mortality rates, changes in fruit quality over time, and switches to control the initiation and termination of diapause. Several equations including fecundity (Eqn 1), development (Eqn 6) and mortality (Eqn 7) rates were fit to experimental data. A similar process was followed for equations required for fruit quality effect on development (Eqns 10, 11 and 13) and diapause processes (Eqn 2). Each of these equations are appropriate in shape for the processes they describe and due to the fitting process, each has several dimensionless shape parameters. A choice to faithfully represent all of *D. suzukii*'s life stages automatically means that all of the life history parameters become stage specific, and a multitude of similar parameters become necessary. Nevertheless, the basic structure of the model is fairly consistent across life stages. Increased pool sizes result from birth or development, and decreased pool sizes result from death or development. The model is, in fact, fairly simple in structure, if somewhat repetitive.

As parsimony is often considered a criterion for deciding between two competing models, I consider an alternate and simplified version of the *D. suzukii* population model. One way to reduce the number of parameters and equations would be to reduce the number of life stages that are explicitly represented. There are two obvious candidates for such an aggregation; larval instars and adult females. *D. suzukii* matures through three larval instar stages which have been represented in the model even though data for the specific temperature-driven development was not collected for these instars. Instead they are represented using the egg-to-pupal rate of development and instar-specific development multipliers were estimated based on the number of days spent in each stage. This approach produces more equations and a repetition of development and mortality parameters. Nevertheless, it can be shown that splitting the egg-to-pupal stage into instars affects the population dynamics. In order to illustrate this, two changes were made in how the model processes juveniles. In one case, I reduced the number of instar stages to one and set the maximum development rate to the sum of the three previous instar values ( $m_{i1} = 0.426$ ). In the second case, I maintained three instar stages but set each of the development rates to be one third the value of used for the single instar version of the model ( $m_{i1-3} = 0.142$ ). This removes any confounding effects of rate variation between the stages. I then ran the model with 100 eggs introduced at time step zero, using a constant temperature profile with diapause and fruit quality ignored and both fecundity and mortality set to zero to more clearly see how the model structure affects stage transit time. The resultant pupae populations for each case are shown in Figure 4.3.

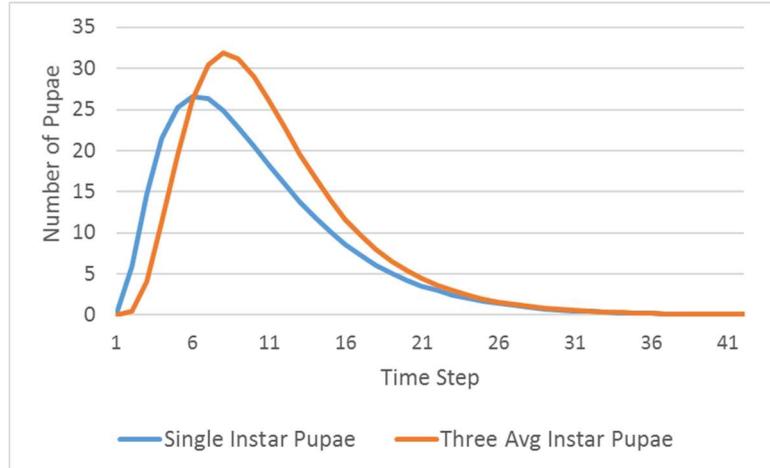


Figure 4.3 Two reconfigurations of the base population model presented in Chapter 2. The first reconfiguration aggregates the three larval instars into a single while the second maintains three instars but averages the entire maximum development rate across all three. For both cases fruit quality and diapaus are ignored, mortality and fecundity are set to 0 and a constant temperature of 23.5°C was used. Resulting pupal population is shown.

In both cases the rate of transition from egg to instar stage remains unchanged but when only a single instar is modelled a more immediate transition to pupa begins due to the higher maximum development rate. This earlier pupal start date results in an earlier transition date from pupa to adult stages as well, and lower peak pupae population occurs as they are “rushed” through the single instar stage. When the maximum development rate is divided across three instars, a delay is seen before the first pupae begin to accumulate. The transition between the three stages creates a time lag, and the pupae population peaks higher and later. What is illustrated here is a simplified case, but the two approaches are impacted by mortality, fecundity rates and perhaps most importantly temperature deviations.

I extend this idea further by aggregating the seven female stages such that the resulting alternate model consists of only five stages (egg, aggregated instar, pupa, male, aggregated female) rather than thirteen. The development rate for the single instar stage ( $m_{i1} = 0.426$ ) is the sum of the development rates for the previously individual instars. Mortality for the single instar and female stages in the simplified model are computed as weighted averages ( $\mu_{i1} = 0.137$  and  $\mu_{F1} = 0.175$ ). In order to isolate the effects of this aggregation on transit dynamics, the flow of individuals back into the model is removed by setting egg viability to zero. Figure 4.4 shows the behaviour of the original and simplified versions of the model for several different temperature profiles previously discussed in Chapter 2. Wayne, MS and Chicoutimi, were selected as they respectively represent the largest and smallest populations produced by the base model for location-specific temperature profiles (see

Figure 2.11). Also, a constant  $23.5^{\circ}\text{C}$  temperature profile was included for comparison. Start parameters include 10000 eggs introduced at time step zero, the diapause sub model is active but the fruit quality sub model is inactive.

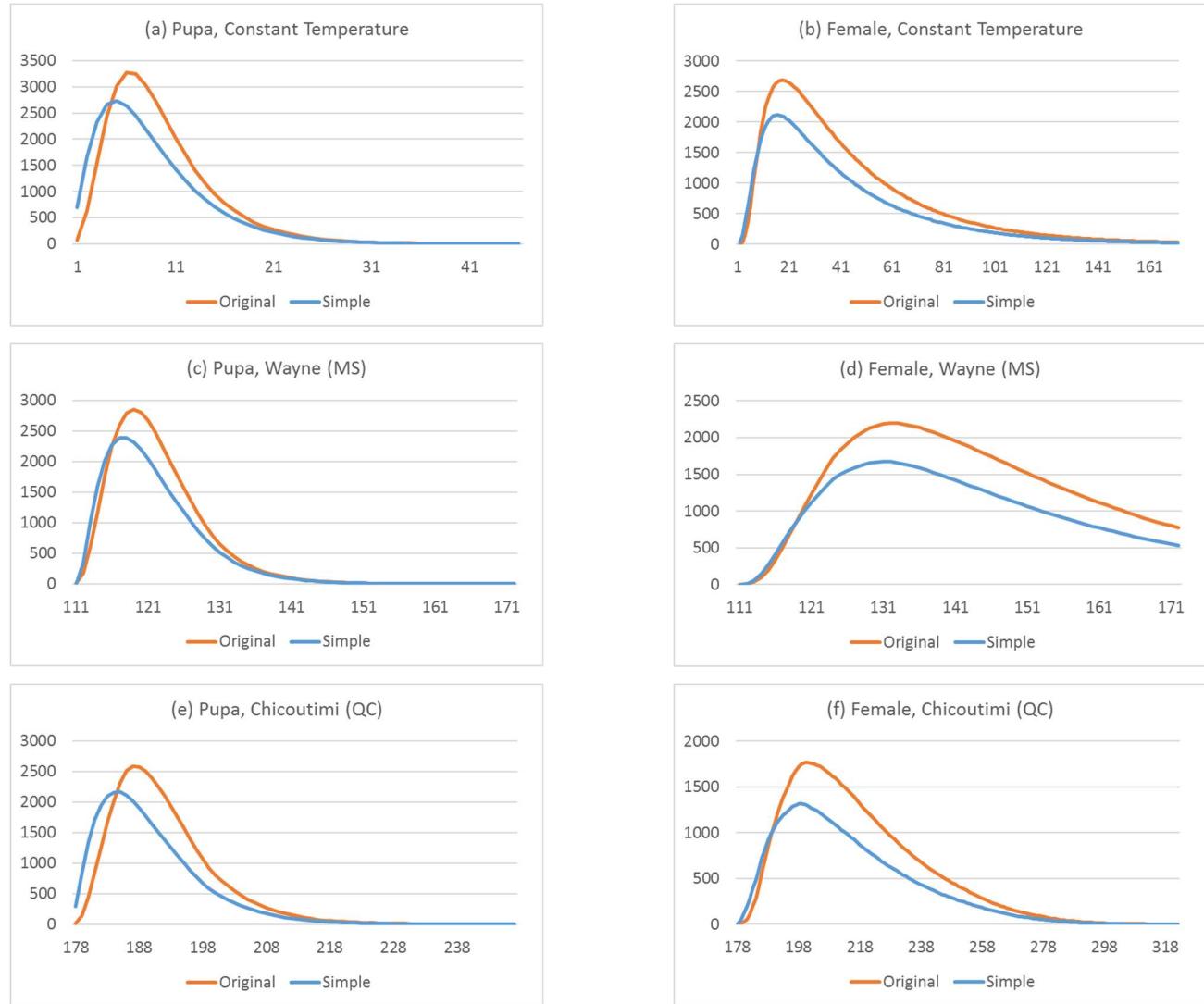


Figure 4.4 Pupal and adult female life stage populations produced by the original and simplified version of the model for various temperature profiles. All graphs have been adjusted at the y-axis begin to highlight productive areas population change. All simulations begin with 10000 eggs at time step zero and include the diapause sub model. Fruit quality is not included.

In each case, the behaviour is similar to the test-case shown in Figure 4.3. The simplified model has a higher rate of transition between instar and pupal stages in the early time steps and a lower peak pupal population. A similar, but somewhat less pronounced effect, is seen in adult female stage. There is a difference in both the peak number of individuals and in the time when the peak

occurs between the two versions of the model with the original (complex) model producing the higher and later peak in all cases.

Given that most of the results presented in previous chapters have focused on total population and not on individual life stages the total population for the same three temperature profiles is shown in Figure 4.5. While the general dynamics of the two models are very similar, the original model produces consistently larger populations.

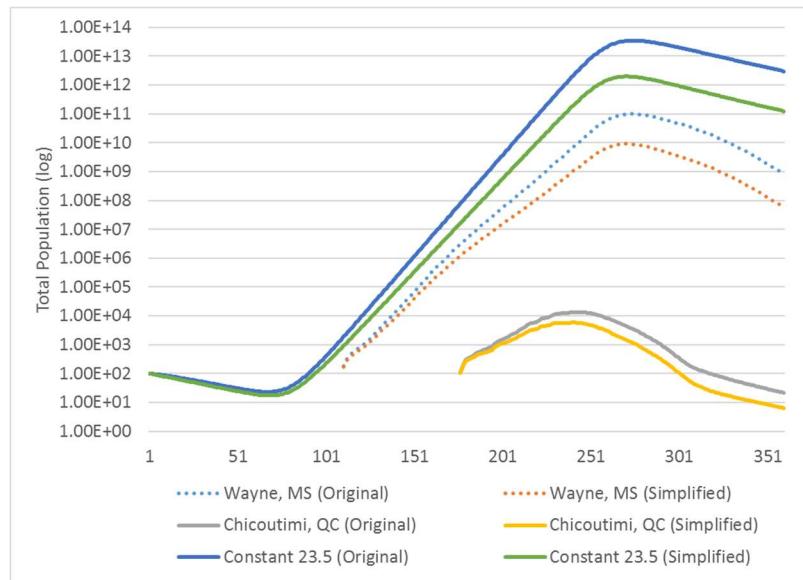


Figure 4.5 Total population (log-scale) for both versions of the model for the same temperature profiles used in the previous figure.

Larger total populations and larger populations for each life stage is a trait of the original model as shown in Figure 4.6 (a) which illustrates the percent difference between the models for various temperature profiles (location-specific averages as well as calculated profiles). The difference in population levels appears to be driven by temperature as the mean temperature curve (b) is very similar to the percent difference curves (a). This is reasonable as temperature is the driving input for both models and the differences between the aggregated and separated life stages is amplified. What is also worth noting is that profiles with larger variation in temperature produce a larger variation in the percent difference between life stages for a given profile. This may be a more important consideration between the two versions of the model; the aggregated life stages with their combined parameters will respond differently to daily variation in temperature than the separate life stages with their specific parameters. The larger the variation in daily temperature difference, the larger the

difference between the two modelled populations. Figure 4.6 (c) highlights the range of difference in peak population dates for the life stages. In general, the peak dates can vary by several days between the two versions of the model. It is likely that other temperature profiles, beyond those tested here, may produce even larger differences in peak population dates. These differences may be significant when considering crops that have a short time period during which they are suitable or preferred food sources for *D. suzukii*.



Figure 4.6 a) Proportional difference between the original and simplified versions of the model for various temperature profiles. b) Mean annual temperature for each of the profiles. Bars indicate standard deviation. c) Mean difference in days (x) including quartiles, standard deviation (bars) and outliers (circles) between the simplified model and the original.

It is difficult to know a priori how different population dynamics will be when all life stages are considered compared to a simplified representation. The intention in developing the model was to illustrate the mechanistic development, mortality and fecundity processes for which data is either known or could reasonably be estimated. While it appears as though a simplified version could have produced qualitatively similar results to the original, the complete set of biologically relevant processes and parameters remain available for experimentation and exploration. To the best of my knowledge, no large scale empirical studies of *D. suzukii* populations have yet been conducted in North America making it difficult to perform a quantitative or qualitative parsimony analysis or to determine which of the two models may be more accurate. While this type of comparison may be feasible as data become available and as more models are constructed, I remind readers that the current objectives of the model are explanation of population-driving processes and exploration of potential spatial patterns due to changes in environmental conditions.

#### 4.3.5 Initialization, Input Data and Sub models

The remaining ODD points are only briefly summarized here for the sake of completeness. Further detail on each of these points can be found in Chapter 2.

Each simulation has a significant number of parameters (Appendix 1) to support the processes of development, mortality and fecundity. Many of these parameters are duplicated between life stages and in most cases, they do not change between simulations. As a result, the model is designed to read these from file. The remaining input for a simulation consists of 7 arguments that may change between simulations. These arguments include the name and path of three files which contain the model process parameters, the temperature data and the output destination, two flags to determine if diapause and fruit sub models are included, the starting number of females and the date at which flies are introduced into the system. An final argument is a legacy feature of the model that was used during development and debugging and tells the model to use hard-coded parameters instead of reading them from file. All of these input arguments and data file formats are described in greater detail below (Executing the Model).

The base model includes all of the key processes (development, mortality, fecundity) for *D. suzukii* population development. Two optional sub models are also available to explore other biologically relevant population drivers including reproductive diapause and quality of available fruit.

The diapause sub model affects female fecundity by activating reproductive processes when the temperature is adequately warm and disabling them when the number of daylight hours becomes too low. The fruit quality sub model affects development and mortality of all life stages to simulate the quality and available of food sources. Further detail on both of these sub models can be found in Chapter 2.

## 4.4 Model Technical Details

The following sections are not part of the ODD protocol but are the technical details to assist those who would download, execute and modify the model.

Although an early Java-language prototype was developed for testing purposes, all of the results presented herein were generated using a C++ version of the model. C++ provides faster execution and a finer level of memory management over Java, both of which were important for the large number of simulations required. C++ libraries and binaries can vary between compilers so I note here that the binaries were compiled using GNU gcc version 4.8.4 for Ubuntu Linux (14.04.03 LTS). Model source code files including headers are freely available at <https://github.com/alangillGuelph/dsPopSim> in order to foster reproducibility, peer-review and potential future extension by third parties the source code. To further support open and free use the software is licensed using a 3-clause Open BSD style license (Morin et al. 2012). The initial release includes compile scripts, several sample temperature files and a sample parameter configuration file. All files include a summary of the files included with the initial release of the model software is included in Table 4.1.

### 4.4.1 Compiling the Models

The model source code makes use of several standard C++ libraries including vector, string, sstream, iostream, fstream, pthread (multi-cell model only), limits, cmath and map. These libraries, while typically made available with most C++ distributions will need to be present before compiling will be

successful. Two scripts named “compileSingleSim.sh” and “compileMultiSim.sh” have been made available with the source code to facilitate compiling. Executing these scripts within the source code directory will produce binary executables, “singleSim” and “multiSim” provided all dependency requirements have been met. singleSim is the base model (see Chapter 2) and will execute for a single population/location/parameter set combination. multiSim is the expanded multi-location version that is explored in Chapter 3. multiSim makes use of the same code used by singleSim to generate the individual populations but manages the location-specific parameters and implements threading to parallelize the execution of multiple locations. There is also a compile script, “compileMultiDemo.sh”, that creates a binary, “multiSimDemo”. This is a simplified version of multiSim that has a hard-coded directory structure for a 3x2 system of locations and is meant as a “toy” version for exploration and testing before working with the larger version.

Files or Directories	Description
Daylight.cpp, EulersMethod.cpp, MultiCellRCP.cpp, MultiCellRCP_demo.cpp, Parameters.cpp, SingleCellRunner.cpp, SolveParameters.cpp, SWDCellMulti.cpp, SWDCellSingle.cpp, SWDPopulation.cpp, SWDSimulatorMulti.cpp, SWDSimulatorSingle.cpp UtilityMethods.cpp	Simulation source code files.
Daylight.h, EulersMethod.h, Parameters.h, SolveParameters.h, SWDCellMulti.h, SWDCellSingle.h, SWDPopulation.h, SWDSimulatorMulti.h, SWDSimulatorSingle.h, UtilityMethods.h	Simulation source code header files.
clark_2003.txt, constantTemperature.txt, tempA_sin.txt, tempB_sin.txt, tempC_sin.txt, tempD_sin.txt	Sample temperature files including Clark County Washington average temperatures (2003), constant temperatures and profiles matching the simulated temperatures described in Chapter 2 (Fig. XXX)
configParams.txt	Sample parameters similar or matching those used in Chapter 2 and 3.
compileMultiSimDemo.sh, compileMultiSim.sh, compileSingleSim.sh	Compile scripts to facilitate creation of binary executables.
TemperatureFiles	Directory with sample temperature files for use with multiSim and multiSimDemo.
DATA	Directory to store sample output files generated by multiSimDemo.

Table 4.1 Summary of files and directories present in the initial source code release.

```

aaron@NOSTROMO:/mnt/c/Users/aaron/Downloads/dsPopSim$ ./compileSingleSim.sh
aaron@NOSTROMO:/mnt/c/Users/aaron/Downloads/dsPopSim$ ./compileMultiSim.sh
aaron@NOSTROMO:/mnt/c/Users/aaron/Downloads/dsPopSim$ ./compileMultiSimDemo.sh
aaron@NOSTROMO:/mnt/c/Users/aaron/Downloads/dsPopSim$ ls -tr
compileMultiSimDemo.sh  SolveParameters.cpp  tempA_sin.txt
clark_2003.txt          SingleCellRunner.cpp SWDSimulatorSingle.h
DATA                    Parameters.h          SWDSimulatorSingle.cpp
constantTemp.txt         Parameters.cpp       SWDSimulatorMulti.h
configParams.txt         SWDCellSingle.cpp  UtilityMethods.cpp
compileSingleSim.sh      SWDCellMulti.h    TemperatureFiles
compileMultiSim.sh       SWDCellMulti.cpp  tempD_sin.txt
EulersMethod.h           SolveParameters.h  tempC_sin.txt
EulersMethod.cpp         SWDSimulatorMulti.cpp UtilityMethods.h
Daylight.h               SWDPopulation.h   singleSim
Daylight.cpp              SWDPopulation.cpp  multisim
MultiCellRCP_demo.cpp   SWDCellSingle.h  multisimDemo
MultiCellRCP.cpp         tempB_sin.txt
aaron@NOSTROMO:/mnt/c/Users/aaron/Downloads/dsPopSim$
```

Figure 4.7 Execution of all three compile scripts within the source directory. The final binaries are shown at the bottom of the right-most column. Example is shown using Bash shell for the Windows 10 Subsystem for Linux including the GNU 4.8.4-2ubuntu1~14.04.3 Ubuntu C++ compiler.

#### 4.4.2 Executing the Models

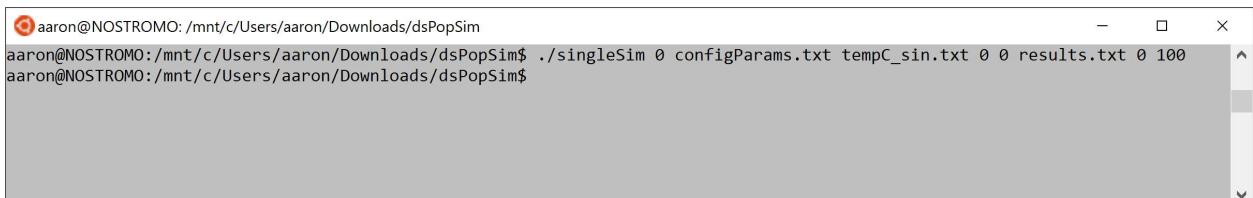
Both of the principle models, singleSim and multiSim have different execution requirements which are explained below.

##### *Single Simulation Model*

For users wishing to execute the simulator for a single population, singleSim requires several command line arguments listed and described in table 4.2. These arguments must appear in the correct order and be of the correct type for execution to be successful. A sample execution is shown in Figure 4.8.

#	Name	Values	Description
1	useDefault	0 or 1	This is legacy functionality and will likely be removed in future versions. The model has several key parameter values hard-coded in case the user fails to provide a complete parameter file. Assuming a complete parameter file is provided the value here should always be 0.  0 = false, do no use defaults 1 = true, use defaults
2	paramFileName	Valid input filename	A valid filename and path (optional) to the parameters required for model execution.
3	tempFileName	Valid input filename	A valid filename and path (optional) to the temperature values required for model execution.
4	ignoreFruit	0 or 1	Determines whether or not the fruit quality sub model will be included during execution. 0 = false, use fruit sub model 1 = true, do not use fruit sub model
5	ignoreDiapause	0 or 1	Determines whether or not the diapause quality sub model will be included during execution. 0 = false, use diapause sub model 1 = true, do not use diapause sub model
6	outputFileName	Valid output file name	A valid output filename and path (optional) for the model results.
7	addFliesDate	-1 or an integer value from 0-364	The timestep (or date ranging from 0 to 364) at which individuals will be included in the simulation. This allows for timesteps to be ignored before introducing the population. -1 is used when the individuals should be introduced at the diapause termination date (temperature-specific, see Chapter 2).
8	startingFemPop	Positive integer value	The initial number of fecund females to introduce into the simulations

Table 4.2 Parameter names, values and descriptions required for the singleSim executable.

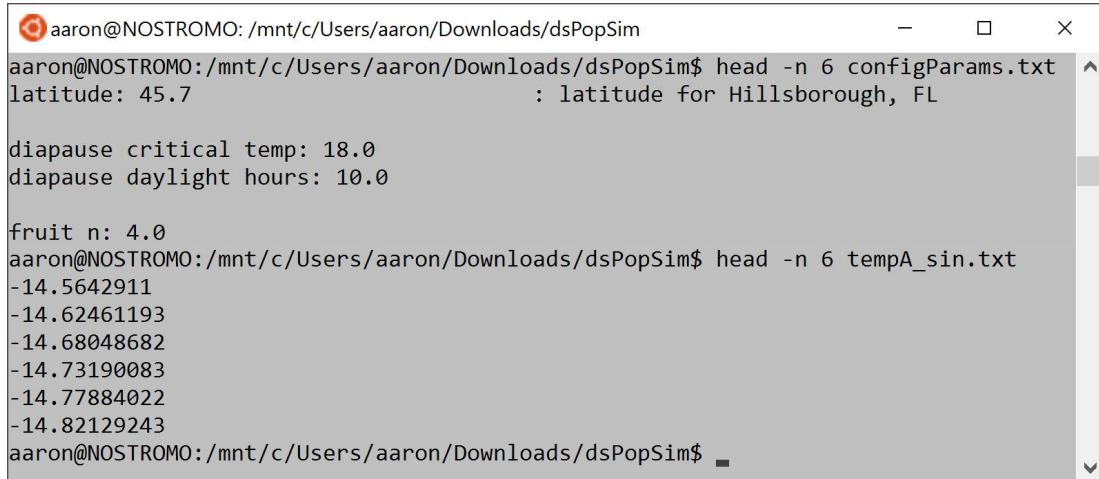


```
aaron@NOSTROMO: /mnt/c/Users/aaron/Downloads/dsPopSim
aaron@NOSTROMO:/mnt/c/Users/aaron/Downloads/dsPopSim$ ./singleSim 0 configParams.txt tempC_sin.txt 0 0 results.txt 0 100
aaron@NOSTROMO:/mnt/c/Users/aaron/Downloads/dsPopSim$
```

Figure 4.8 Sample execution of singleSim using the parameters described in Table 4.2.

While the names of all input and output files are left to users' discretion, input files must be properly formatted in order to be read by the simulator. All temperature files should contain 365 one-per-line entries where each may be integer or floating-point. No other information should appear in this file. Parameter configuration files should include colon (:) separated key-value pairs, one-pair-per-line. This file can include comments and user-specific information in addition to the key-value pairs but care should be taken not to complicate the file unnecessarily given the number of

parameters required to successfully execute the simulation. The first five values for sample temperature and parameter configuration files are show in Figure 4.9. For a list of all of the configurable parameters please see Appendix 1 and the configParams.txt file included with the initial release.



```
aaron@NOSTROMO: /mnt/c/Users/aaron/Downloads/dsPopSim
aaron@NOSTROMO:/mnt/c/Users/aaron/Downloads/dsPopSim$ head -n 6 configParams.txt
latitude: 45.7                                : latitude for Hillsborough, FL

diapause critical temp: 18.0
diapause daylight hours: 10.0

fruit n: 4.0
aaron@NOSTROMO:/mnt/c/Users/aaron/Downloads/dsPopSim$ head -n 6 tempA_sin.txt
-14.5642911
-14.62461193
-14.68048682
-14.73190083
-14.77884022
-14.82129243
aaron@NOSTROMO:/mnt/c/Users/aaron/Downloads/dsPopSim$
```

Figure 4.9 The first five entries for samples of both a temperature file and a parameter configuration file included with the initial release of the model.

Upon successful execution, the results will be stored in the user-defined output file. This file has tab-delimited columns of data for each of the life stages (female stages aggregated) and the fruit quality index. This format can easily be imported into post-processing software such as Matlab or Excel. A summary section is included at the bottom of the file. Cumulative and peak populations for each stage as well as the date (time step) at which the peak population occurred. Also included is the diapause injection date, for reference and the date at which peak fruit quality was achieved. A sample result data file is shown in Figure 4.10.

```

aaron@NOSTROMO:/mnt/c/Users/aaron/Downloads/dsPopSim
aaron@NOSTROMO:/mnt/c/Users/aaron/Downloads/dsPopSim$ ./singleSim 0 configParams.txt tempC_sin.txt 0 0 results.txt 0 100
aaron@NOSTROMO:/mnt/c/Users/aaron/Downloads/dsPopSim$ head -10 results.txt      t tempC_sin.tx
Time:   eggs:    instar1:    instar2:    instar3:    pupae:    males:    females:    fruit:
0       0       0       0       0       0       99.8341  0.0500807
1       0       0       0       0       0       96.5525  0.0501613
2       0       0       0       0       0       93.3254  0.0518021
3       0       0       0       0       0       90.1488  0.0534938
4       0       0       0       0       0       87.0219  0.0552382
5       0       0       0       0       0       83.9464  0.0570368
6       0       0       0       0       0       80.9251  0.0588917
7       0       0       0       0       0       77.9613  0.0608046
8       0       0       0       0       0       75.0586  0.0627776
aaron@NOSTROMO:/mnt/c/Users/aaron/Downloads/dsPopSim$ tail -n 20 results.txt
357     13.6249 7.79471 11.6214 26.4124 580.899 579.018 2627.06 0.05
358     12.826 7.31833 10.9442 24.9357 555.911 549.216 2529.04 0.05
359     12.0716 6.86905 10.3026 23.5358 532.038 520.835 2434.07 0.05

Total Cumulative Populations
2.64805e+06 1.47975e+06 1.64927e+06 2.2973e+06 3.48161e+06 1.41032e+06 2.47786e+06

Peak Populations
27671.7 15319.5 16649.6 21848.7 28038.2 9254.57 15849.9

Peak Populations Day
173.95 173.05 175.05 177.95 186 277 281.2

Diapause injection date: 0

Fruit max date: 87
aaron@NOSTROMO:/mnt/c/Users/aaron/Downloads/dsPopSim$
```

Figure 4.10 Output showing the first 10 lines and the last 20 lines of a sample simulation.

#### 4.4.3 Multiple Simulation Model

In order to facilitate the production of the many data sets required for the climate change exploration presented in Chapter 3, a new execution strategy was developed. The multiSim model uses the same population algorithms and source code but is set up to more effectively handle the parameter and output requirements of simulating multiple locations. At the time of the initial release of the source code and required files the multiSim model uses a hard-coded directory structure. During execution, it reads from and outputs result files to these directories. Every file in the input directory is processed, and assuming successful execution, a corresponding output file (as discussed previously) is produced.

Input files for the multiSim model are stored in a separate directory. As an example, the for the multiSimDemo executable these files are store in a directory called “TemperatureFiles”. Each file in this directory contains a temperature file formatted as described above. The naming convention of these files (eg. randomTemp\_lat2\_lon1.txt) specifies the relative latitude and longitude. These relative coordinates are parsed during execution are converted to absolute coordinates. Absolute

latitude is required by the base model for determining diapause termination while longitude is required only to reassemble the results into a two-dimensional representation post-execution. Output files are written to a different directory (“DATA” in the case of multiSimDemo) and for each input temperature file two files separate files containing the per-time-step data and the summary data are created. This separation facilitates post-processing as the summary data may be handled differently than the per-time-step results. The remaining parameters required to run the base simulation within the multiSim executor are again hard-coded but can be modified as needed. Any modifications require a recompiling of the code before execution. A sample of the lines of code that can be altered in the MultiSimRCP\_demo.cpp source file are described in Table 4.3 while an example execute and view of the directory structure are shown in Figure 4.11.

Code to be altered	Description
int rows = 3; // latitude int cols = 2; // longitude	This specified the number of locations to be processed as a number of latitude x longitude cells.
bool ignoreFruit = false; bool ignoreDiapause = false; int startDay = -1;	These are analogous to the single simulator parameters discussed previously and can be changed as needed. All cells in the multiSim execution will use these parameters.
double lat = 24.5;	This is the starting latitude and should be adjusted for the locations being simulated.
"randomTemp_lat" << (i + 1) << "_lon" << (j + 1) << ".txt";	This line assumes that the input temperature file name will be randomTemp_lat#_lon#.txt and can be adjusted as needed. However, the value used in place of # must match the loop increment values.
"DATA/output_" "DATA/summary_"	Output files will be written to subdirectory DATA within the current directory. Per-time step data and summary data file will be prepended with output_ and summary_ respectively.
"TemperatureFiles/"	Input files will be read from subdirectory TemperatureFiles within the current directory.
SWDSimulatorMulti sim("", rows, cols, latitudes);	“” indicates the use of the hard-coded parameters but can be changed to use any parameters files. Eg. “configParams.txt” would use a parameters file called configParams.txt located in the current directory.

Table 4.3 Code that can be altered to customize the multiSim or multiSimDemo execution. Any code modifications require a re-compile before execution.

```

aaron@NOSTROMO:/mnt/c/Users/aaron/Downloads/dsPopSim$ ls TemperatureFiles/
randomTemp_lat1_lon1.txt randomTemp_lat2_lon1.txt randomTemp_lat3_lon1.txt
randomTemp_lat1_lon2.txt randomTemp_lat2_lon2.txt randomTemp_lat3_lon2.txt
aaron@NOSTROMO:/mnt/c/Users/aaron/Downloads/dsPopSim$ ls DATA/
aaron@NOSTROMO:/mnt/c/Users/aaron/Downloads/dsPopSim$ ./multiSimDemo
aaron@NOSTROMO:/mnt/c/Users/aaron/Downloads/dsPopSim$ ls DATA/
output_randomTemp_lat1_lon1.txt summary_randomTemp_lat1_lon1.txt
output_randomTemp_lat1_lon2.txt summary_randomTemp_lat1_lon2.txt
output_randomTemp_lat2_lon1.txt summary_randomTemp_lat2_lon1.txt
output_randomTemp_lat2_lon2.txt summary_randomTemp_lat2_lon2.txt
output_randomTemp_lat3_lon1.txt summary_randomTemp_lat3_lon1.txt
output_randomTemp_lat3_lon2.txt summary_randomTemp_lat3_lon2.txt
aaron@NOSTROMO:/mnt/c/Users/aaron/Downloads/dsPopSim$
```

Figure 4.11 Directory structure and contents before and after execution of multiSimDemo (similar for multiSim). Directory TemperatureFiles contains properly formatted input temperature files with a naming convention that supports iterative processing. Corresponding per-time step and summary result files are created in output directory DATA.

#### 4.4.4 Web-based Model

The command-line version of the single and multi-cell models offer researchers a high level of flexibility for both parameterization and output but there can be significant post-processing required depending on the data goals. In order support a holistic approach to *D. suzukii* management and mitigation (Beers et al. 2011) by reaching a more general audience including policy-makers, berry producers and pest species enthusiasts, a web-based interface for the single-location simulator (<https://leecs.cs.laurentian.ca/>) was created with both simpler input requirements and output options.

The web-based version allows users to select from various uploaded temperature profiles or to upload their own files. Per the simulator requirements, temperature files must be plain text with a single temperature entry per line. For each graph panel, the user may select different parameters including temperature file, number of female flies to be introductory, day of introduction of females and whether or not to ignore the diapause and fruit sub models. See Figure 4.12 for an illustration of the configuration panel. Users can independently configure each of the four panels (Figure 4.13) in order to compare the output of different configurations. For simplicity of output instar life stages are aggregated as are adult female stages.

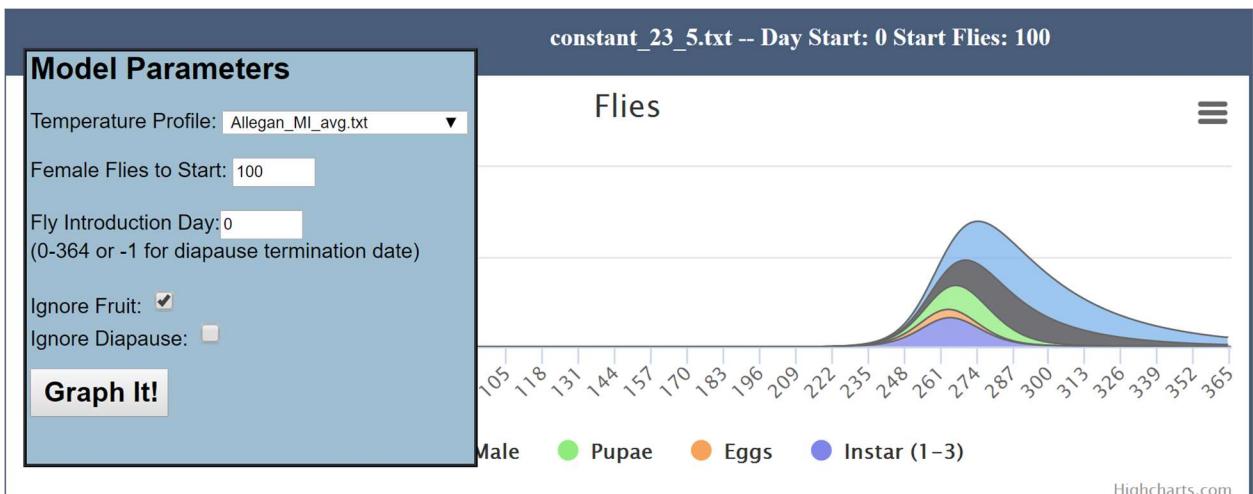


Figure 4.12 Configuration options for each of the graph panels of the web-based simulator. Each panel can be configured independently.



Figure 4.13 Four independently configured graph panels.

The graphs in the web-based version use Chart.js, an open source graphing platform that allows users to hover over the graph to see inspect data points and the ability to print or save the graphs in a variety of formats.

## 4.5 Discussion

Models are powerful and increasingly popular tools in ecological research but their utility is greatly diminished without effective communication of their design, implementation and use. In this chapter I presented a modified version of the ODD protocol (Grimm 2006, 2010) to augment the model description in Chapter 2. A discussion on parsimony was included in the model description. Although I feel that the *D. suzukii* model objectives of illustrating the biological mechanisms of population dynamics warrant its mechanistically-rich design, a simpler model was developed for discussion purposes. The alternate model produced qualitatively similar results though it's "explanatory power" (Laurenroth 2003) is somewhat reduced. I provide a "user manual" that describes how obtain, compile and execute the model as well as a description of the input data formats in order to foster result reproduction, peer-review and future extension by third parties. Finally, a brief overview of the simplified web-based version of the model is presented. This version of the model was developed as a more accessible tool for use by a broader audience of *D. suzukii* stakeholders (Beers et al. 2011).

## 5 General Discussion and Conclusion

### 5.1 Introduction

The principle objective of this work was to develop a mechanistic model of *D. suzukii* population dynamics and to use it to investigate the possible impacts of climate change on this species. Through this exercise several modelling objectives (see Chapter 1) were satisfied. Perhaps the most important of these objectives is explanation and exploration. Choosing a mechanistically-rich approach allows the model to illustrate the effects of temperature on *D. suzukii* population development and mortality and these effects can be isolated at each life stage and for the temperature-dependent sub models of diapause and fruit quality. By parameterizing the model for different locations, I illustrate the temperature-dependent differences (both contemporary and considering future climatic change) in infestation potential. Perhaps more importantly, I have created a tool, and implemented in an environment accessible to all, that others may use to ask many other questions about this species and the risks associated with it. Given the economic, environmental and human costs of management strategies gauging infestation potential is a key aspect in determining appropriate location-specific solutions. In the following sections, I provide specific closing remarks for each of the principle chapters (2-4) of this work.

### 5.2 Mechanistically Modelling *D. Suzukii* Population Dynamics

Despite the amount of available data on development, mortality and fecundity processes of *D. suzukii*, there remain several knowledge gaps that the presented model helped to highlight. This is, after all, one of the purposes for constructing a mechanistic model — gap identification. In this section I address these gaps including specifics of extrinsic mortality, fruit host quality on development and mortality and finally overwintering. Also, included in this section is a brief discussion on the need for further study on *D. suzukii* migration mechanisms, field trap techniques for quantifying population levels in the wild and empirical estimates of commercial crop damage and loss. I conclude with a brief

commentary on model validation.

### 5.2.1 Identifying *D. Suzukii* Knowledge Gaps

Chapter 2 acknowledges the importance of extrinsic mortality on population size. It is well known that natural enemies are an important control on pest populations. For example:

“In a year aphids, could form a layer 149 km deep over the surface of the earth. Thank God for limited resources and natural enemies.” (Richard Harrington, 1994 as cited in Leather, date unknown)

In terms of incorporating extrinsic mortality into the model, for now it remains only a “placeholder” rate variable. As data continues to accrue on the particular mechanisms, regionality and quantification of extrinsic mortality, this placeholder can be replaced with a more biologically-relevant value or sub-process. Extrinsic mortality factors including natural predators, parasites, pathogens, etc. Recent research into natural control factors for *D. suzukii* suggests that viral (Cini et al. 2012) and parasite threats including parasitoid wasps (Kacsoh and Schlenke 2012, Poyet et al. 2013) nematodes and fungi exist (Woltz et al. 2015), but that their efficacy at limiting population sizes is limited. Many of the previously cited studies were carried out using predators that are, like *D. suzukii*, native to Japan and, to date, there is little information on predator or pathogen efficacy in North America. Furthermore, existing studies are laboratory-based and extrinsic mortality in the wild will likely prove even more difficult to quantify. Similarly, further quantification of development rates on specific commercial fruit (Tochen et al. 2014) and wild hosts (Lee et al. 2015, Chapman et al. 2016) would allow for a more mechanistic fruit model to be developed. As with extrinsic mortality, this sub model could be made regionally-specific and could account for multiple host types and harvesting periods with the appropriate data. Finally, while reproductive diapause is presumed for *D. suzukii* (Dalton 2010) the specific thresholds and mechanistic details for induction and termination for this species remain unknown. While I feel that the first-pass approximation of diapause based on similar species is adequate for illustration of this important process, as more data become available, these parameter estimates can be refined.

Overwintering of populations is a complex topic that requires attention beyond the scope of the work presented here. While data exists on lower limits of temperature where mortality occurs and development ceases, research is ongoing to determine freeze periods (duration, temperature limits and life stage specificity) from which *D. suzukii* can recover (Oshtu et al. 1993). Similarly, there is research to suggest that inter-generational cold hardiness and plasticity may also play a role in cold tolerance (Jacobs et al. 2015, Wallingford et al. 2016, Wallingford and Loeb 2016, Shearer 2016, Kimura et al. 2004). Finally, overwintering habitats, both anthropogenic and in the wild are subjects of ongoing study (Stephens et al. 2015). Given these complexities, I decided to exclude overwintering for the initial version of the model.

The current version of the model is spatially implicit such that each location-specific parameterization is independent of any others. At this time, there is no account of immigration or emigration. Although it is difficult to predict the exact effects this may have compared to the presented results it would add a layer of realism that is not currently present. Short term anthropogenic events (such as accidental transportation) could be a significant mechanism of introduction and range expansion. Longer term environmental phenomena such as drought, changes in prevailing winds, shifting of preferred host range could also have an impact on *D. suzukii* migration patterns and alter localized infestation potential over time. While more data is required on the specifics of *D. suzukii* migration there is precedent for such spatially explicit insect-pest modelling including (Storer 2003, Prasad et al. 2010, Vinatier et al. 2009).

The final data gaps address here are those of field trap data and commercial loss estimates. Among other data, field traps help to identify changes in range and can offer insight into wild population levels. This type of empirical data is often used to validate models however the only attempt to do so with a mechanistic model to date (Wiman et al. 2014) found current trap data to be insufficient for validation. Specifically, they highlighted trap placement, trap set setup and a relatively small ratio of adults (found in traps) to juvenile stages (fruit-damaging) among reasons to why trap counts are a poor indicator of infestation potential. They propose experimental changes for improved trap data as well as controlled experiments such as caged populations as alternatives for model validation. As higher quality trap data can only help to improve model calibration and validation efforts I echo the call for continued empirical research on quantifying wild population levels. Finally, I am aware of only one study attempting to quantify economic loss (Bolda 2010) which bases estimates on theoretical infestation losses of 20%. While this provides a baseline figure for various fruit crop types empirical estimates of both product loss and mitigation costs can be correlated with

modelled infestation potential to provide a more holistic picture for management strategizing.

### 5.2.2 Model validation

To this point little has been said of model validation though it is often considered an important part of the modelling process. One of the most commonly applied methods of validation is comparison of model results with empirical data (discussed previously with the model of Wiman et al. 2014, Power 1993, Mayer and Butler 1999, Legates and McCabe 1999). As with degree of parsimony, this type of model validation may be useful when selecting between competing models. However, this represents just one type of model validation technique. Thornley and Johnson (1990) advise against rigidly defining successful validation only in quantifiable terms, suggesting instead that success of applied models should be considered in conjunction with the modeler's objectives. They further proceed to state that a model's validity can be as simple as an improvement over current models even if only under a limited set of circumstances (see also Hilborn and Mangel, 1997, Grimm and Railsback 2005):

"All that needs to be demonstrated is that the proposed model produces better results than the current model (current practice) in a defined set of circumstances; this is sometimes referred to as the 'champion challenge' approach."

Rather than pigeonholing validation strictly as a measure of the model's ability to reproduce observed data, a more flexible approach considers a model's ability to meet its intended objectives. Thornley (1998) summarizes as follows:

"Validation' is another problem area where confusion abounds. It needs to be said first that the validity of model is a quantitative concept, ranging say between zero and unity, and second that validity is not an attribute of the model itself. Validity depends on the relationship between a model and a set of objectives. A model may be valid for managing grassland in southern England (say) but fail in the wetter west of Britain, and also fail to contribute to our understanding of how process is related to outcome. It is self-evident that a model is valid as seen by the creator of the model, who set the modeling objectives. A model is generally less valid to other scientists, who legitimately evaluate the model with their own objectives in mind.

An important question is: is it reasonable for referees and editors to evaluate a model from the standpoint of their own private objectives, rather than the author's objectives? I think not, but the opposite is all too common."

With this in mind, the model I have presented in this work may be considered an improvement over the currently limited suite of existing models *D. suzukii* population models (Wiman et al. 2014, Gutierrez et al. 2016) in several ways including a more mechanistically-rich design which provides supplemental explanatory potential. It meets its intended objectives as a tool of exploration of population dynamics under varying environmental conditions.

### 5.3 Exploring the Implications of Climatic Change

Climate scientists agree that temperatures will continue to rise over the next century and while much is known about the mechanisms of this change, the scale and complexity of the interactions involved make it difficult simulate even on current super computers. As a result, Global Circulation Models make different assumptions regarding the driving forces behind climatic change and these hybrid mechanistic-phenomenological models produce "plausible views" of future climatic conditions (Newman et al. 2011). The subset of CMIP5 models used in Chapter 3 produce possible temperature outcomes based on the standardized set of Relative Concentration Pathways, which make assumptions based on potential socio-economic factors driving radiative forces. Based on the differences between GCM process and parameter assumptions and the differences between RCP output assumptions there is a significant level of divergence in simulated temperatures for any given location. Unsurprisingly, the range of temperatures and the associated uncertainty increase with time (see Figure 5.1, see also Figure 3.5). Rather than looking at these outputs as predictions, they should be treated as explorations of the potential range of temperatures. Consequently the *D. suzukii* population levels that are illustrated in chapter 3 are not predictions but represent an exploratory range of possible outcomes.

The GCM-RCP consensus results show a mixed future for *D. suzukii* in North America over the next century. In particular, *D. suzukii* appears to be well-suited for the contemporary climate but as temperatures continue to increase population levels drop over much of the continent. This suggests that for many berry-producing locations this pest may be a 'relatively' short-lived concern. The

exceptions are locations to the north including southern Canada which may see an increase in *D. suzukii* population levels as winters shorten, growing seasons lengthen and average temperatures become better suited for sustained development. However, these results do not take into account several confounding issues that are likely to impact the reality of *D. suzukii*'s future.

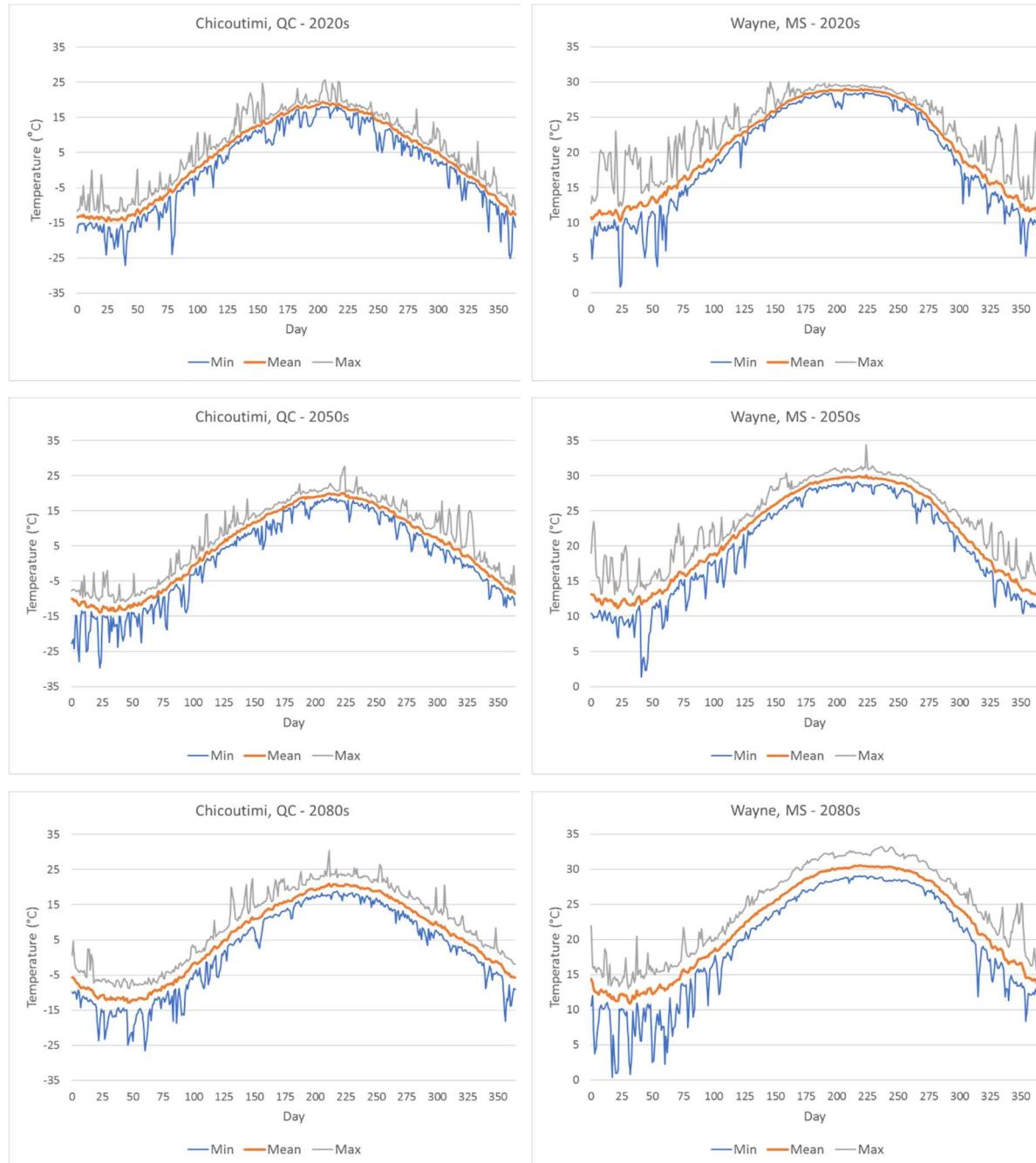


Figure 5.1 Daily maximum, minimum and mean temperatures across all GCM-RCP combinations for Chicoutimi QC and Wayne MS. In general, maxima and minima diverge from the mean with time.

The relationship between environmental suitability for soft-skinned fruit and for *D. suzukii* is not necessarily one-to-one. It is likely that the model results are identifying locations of high infestation potential where there is little or no soft-skinned fruit industry. Conversely some soft-skinned fruit producing locations that are likely on the fringes of *D. suzukii*'s thermal tolerances and while a positive identification of the pest may be made, the potential for significant economic loss may be relatively low. This mismatch of fruit-presence versus *D. suzukii* infestation potential may be further aggravated as climatic change causes shifts in the range of both. Anthropogenic solutions such as large scale green housing which would provide suitable habitats for both fruit and *D. suzukii* where perhaps neither could thrive otherwise are also not considered here.

Assuming the rate of climatic change does not outpace *D. suzukii*'s ability to adapt to increasing temperatures it is reasonable to assume that some variation in thermal tolerance will occur over the coming decades. While there is evidence of inter-generational cold-hardiness and latitudinal clines in both *D. suzukii* and similar species, research is on-going. It would be largely speculative at this time to add an evolutionary sub model based on the limited existing data or one extrapolated from data on similar species although this would be an interesting and worthwhile extension to the model.

In spite of its limitations, the presented model provides a long-term view of *D. suzukii* population potential. Given the inherent uncertainty in the temperature data used to drive the model and the current simplifying assumptions in the model, my goal is not to promote a specific 80-year management decisions. Rather I provide a multi-tiered outlook into contemporary, mid-range and long-term population potential that can be iteratively refined using shorter-term input data, model adjustments and model extensions.

## 5.4 Improvement Through Collaboration and Open Access

By releasing the model source code, providing a detailed overview of functionality and design decisions, and through open licensing I hope to encourage collaboration and extension of this preliminary work. In addition to the provided documentation (see Chapter 4), the model has been designed using standard tools and practices (such as object oriented design) which should help to facilitate third-party development efforts.

By providing a web-based version of the model I hope to promote its use by a larger audience of stakeholders including extension specialists and those responsible for mitigation and management decision-making. Reaching a broad audience is an important part of promoting discussion on the risks and the strategies for managing this potentially destructive pest. The web interface uses the same base population dynamics model (see Chapter 2) with a reduced set of input parameters and a more graphically-rich output format. The multi-window output format allows for the comparison between different parameters sets or temperature profiles. All of the temperature profiles used in Chapter 2 have been uploaded to the web-based model so that parameter variations can be explored. More importantly, users can upload up to ten custom temperature profiles so that they might explore localized climate potential or use the website as their own “what-if” scenario generator. Graphical and text-based population output can be exported for easy archiving or sharing. While the majority of the development, fecundity and mortality parameters have been hidden in the initial web interface, an enhanced interface could be developed to allow access to all of the parameters. Finally, a database or archiving extension could be developed to allow for immediate access to previously configured and executed simulations.

The remainder of Chapter 4 consists of a discussion on the role of parsimony particularly with respect to mechanistically-rich models. While reducing model parameters and complexity is often a worthwhile goal, it must be considered in context of model objectives. Using parsimony as a criterion to select between competing models is often perfectly reasonable. However, if the objective of a model is to illustrate the natural complexity of a system, to explain the specific mechanisms that give rise to a particular phenomenon or to explore various scenarios tied to those specific mechanisms then removing these processes for the sake of parsimony is counter-productive. While an alternate and more parsimonious version of the *D. suzukii* population model was able to produce qualitatively similar results there were some important changes in the model’s behaviour. Most notably the rate through which eggs became adults was affected, peak population dates changed and overall population levels differed between the two versions of the model. In terms of which model is more *numerically* accurate in reproducing observed populations of *D. suzukii*, I cannot say here due to a lack of observed data, but the model as implemented in Chapter 2 is more *biologically* accurate. Similarly, I cannot say whether the model I am presenting is more accurate than the model of Guitierrez et al. (2016) that presents only a single egg-to-adult stage. However, the model I have presented in this work represents the most mechanistically-realistic model of *D. suzukii* population dynamics to date and as such offers the most explanatory power and flexibility for parameterization and scenario

exploration.

## 6 Bibliography

Andreazza F, Haddi K, Oliveira EE, Ferreira JA. *Drosophila suzukii* (Diptera: Drosophilidae) Arrives at Minas Gerais State, a Main Strawberry Production Region in Brazil. *Florida Entomologist*. 2016 Dec;99(4):796-8.

Asplen MK, Anfora G, Biondi A, Choi DS, Chu D, Daane KM, Gibert P, Gutierrez AP, Hoelmer KA, Hutchison WD, Isaacs R. Invasion biology of spotted wing Drosophila (*Drosophila suzukii*): a global perspective and future priorities. *Journal of Pest Science*. 2015 Sep 1;88(3):469-94.

Atallah J, Teixeira L, Salazar R, Zaragoza G, Kopp A. The making of a pest: the evolution of a fruit-penetrating ovipositor in *Drosophila suzukii* and related species. *Proceedings of the Royal Society of London B: Biological Sciences*. 2014 Apr 22;281(1781):20132840.

Austin M. Species distribution models and ecological theory: a critical assessment and some possible new approaches. *Ecological modelling*. 2007 Jan 10;200(1):1-9.

Awkerman JA, Hemmer B, Almario A, Lilavois C, Barron MG, Raimondo S. Spatially explicit assessment of estuarine fish after Deepwater Horizon oil spill: tradeoffs in complexity and parsimony. *Ecological Applications*. 2016 Mar 1.

Bale JS, Masters GJ, Hodkinson ID, Awmack C, Bezemer TM, Brown VK, Butterfield J, Buse A, Coulson JC, Farrar J, Good JE. Herbivory in global climate change research: direct effects of rising temperature on insect herbivores. *Global Change Biology*. 2002 Jan 1;8(1):1-6.

Baumgärtner S, Becker C, Frank K, Müller B, Quaas M. Relating the philosophy and practice of ecological economics: The role of concepts, models, and case studies in inter-and transdisciplinary sustainability research. *Ecological Economics*. 2008 Oct 15;67(3):384-93.

Beers EH, Van Steenwyk RA, Shearer PW, Coates WW, Grant JA. Developing *Drosophila suzukii* management programs for sweet cherry in the western United States. *Pest management science*. 2011 Nov 1;67(11):1386-95.

Bentz BJ, Régnière J, Fettig CJ, Hansen EM, Hayes JL, Hicke JA, Kelsey RG, Negrón JF, Seybold SJ. Climate change and bark beetles of the western United States and Canada: direct and indirect effects. *BioScience*. 2010 Sep 1;60(8):602-13.

Berzitis EA, Minigan JN, Hallett RH, Newman JA. Climate and host plant availability impact the future distribution of the bean leaf beetle (*Cerotoma trifurcata*). *Global change biology*. 2014 Sep 1;20(9):2778-92.

Besheres SN, Fewell JH. Models of division of labor in social insects. *Annual review of entomology*. 2001 Jan;46(1):413-40.

Bolda MP, Goodhue RE, Zalom FG. Spotted wing drosophila: potential economic impact of a newly established pest. *Agricultural and Resource Economics Update*. 2010;13(3):5-8.

Bossel H. Systems and models: complexity, dynamics, evolution, sustainability. BoD—Books on Demand; 2007.

Briere JF, Pracros P, Le Roux AY, Pierre JS. A novel rate model of temperature-dependent development for arthropods. *Environmental Entomology*. 1999 Feb 1;28(1):22-9.

Brekke L, Thrasher BL, Maurer EP, Pruitt T. Dowscaled CMIP3 and CMIP5 climate projections: Release of downscaled CMIP5 climate projections, comparison with preceding information, and summary of user needs. Technical Service Center, Bureau of Reclamation, US Department of the Interior, Denver, CO. 2013;1.

Bruck DJ, Bolda M, Tanigoshi L, Klick J, Kleiber J, DeFrancesco J, Gerdeman B, Spitler H. Laboratory and field comparisons of insecticides to reduce infestation of *Drosophila suzukii* in berry crops. *Pest management science*. 2011 Nov 1;67(11):1375-85.

Burrack HJ, Fernandez GE, Spivey T, Kraus DA. Variation in selection and utilization of host crops in the field and laboratory by *Drosophila suzukii* Matsumara (Diptera: Drosophilidae), an invasive frugivore. *Pest management science*. 2013 Oct 1;69(10):1173-80.

Chornesky EA, Bartuska AM, Aplet GH, Britton KO, Cummings-Carlson J, Davis FW, Eskow J, Gordon DR, Gottschalk KW, Haack RA, Hansen AJ. Science priorities for reducing the threat of invasive species to sustainable forestry. *Bioscience*. 2005 Apr 1;55(4):335-48.

Coop, L. 2010. Online phenology and degree-day model for agricultural and decision-making in the US. Integrated Plant Protection Center, Botany & Plant Pathology Dep. Oregon State University, Corvallis, Oregon. (<http://uspest.org/risk/models?spp=swd>).

Crozier L, Dwyer G. Combining population-dynamic and ecophysiological models to predict climate-induced insect range shifts. *The American Naturalist*. 2006 Jun;167(6):853-66.

Dalton DT, Walton VM, Shearer PW, Walsh DB, Caprile J, Isaacs R. Laboratory survival of *Drosophila suzukii* under simulated winter conditions of the Pacific Northwest and seasonal field trapping in five primary regions of small and stone fruit production in the United States. *Pest management science*. 2011 Nov 1;67(11):1368-74.

Damus M. Some preliminary results from CLIMEX and MAXENT distribution modelling of *Drosophila suzukii*. Version 2. CIFA Plant Health Assessment, Ottawa, Canada. 2009.

Davis MA, Pelsor M. Experimental support for a resource-based mechanistic model of invasibility. *Ecology letters*. 2001 Aug 14;4(5):421-8.

Diepenbrock LM, Rosensteel DO, Hardin JA, Sial AA, Burrack HJ. Season-long programs for control of *Drosophila suzukii* in southeastern US blueberries. *Crop Protection*. 2016 Mar 31;81:76-84.

Dreves AJ. IPM program development for an invasive pest: coordination, outreach and evaluation. *Pest management science*. 2011 Nov 1;67(11):1403-10.

Duarte CM, Amthor JS, DeAngelis DL, Joyce LA, Maranger RJ, Pace ML, Pastor J, Running SW. The limits to models in ecology. *Models in Ecosystem Science*. Princeton University Press, Princeton, New Jersey. 2003:437-51.

Emiljanowicz LM, Ryan GD, Langille A, Newman J. Development, reproductive output and population growth of the fruit fly pest *Drosophila suzukii* (Diptera: Drosophilidae) on artificial diet. *Journal of economic entomology*. 2014 Aug 1;107(4):1392-8.

Estay SA, Lima M, Labra FA. Predicting insect pest status under climate change scenarios: combining experimental data and population dynamics modelling. *Journal of Applied Entomology*. 2009 Aug 1;133(7):491-9.

Evans MR, Grimm V, Johst K, Knuutila T, De Langhe R, Lessells CM, Merz M, O'Malley MA, Orzack SH, Weisberg M, Wilkinson DJ. Do simple models lead to generality in ecology?. *Trends in ecology & evolution*. 2013 Oct 31;28(10):578-83.

Gilchrist GW, Huey RB, Balanyà J, Pascual M, Serra L. A time series of evolution in action: a latitudinal cline in wing size in South American *Drosophila subobscura*. *Evolution*. 2004 Apr;58(4):768-80.

Gillman M. An introduction to mathematical models in ecology and evolution: time and space. John Wiley & Sons; 2009 Apr 8.

Glarner H, (2006, January 15). Length of Day and Twilight. Retrieved from [http://www.gandraxa.com/length\\_of\\_day.xml](http://www.gandraxa.com/length_of_day.xml)

Goodhue RE, Bolda M, Farnsworth D, Williams JC, Zalom FG. Spotted wing drosophila infestation of California strawberries and raspberries: economic analysis of potential revenue losses and control costs. *Pest management science*. 2011 Nov 1;67(11):1396-402.

Grimm V, Berger U, Bastiansen F, Eliassen S, Ginot V, Giske J, Goss-Custard J, Grand T, Heinz SK, Huse G, Huth A. A standard protocol for describing individual-based and agent-based models. *Ecological modelling*. 2006 Sep 15;198(1):115-26.

Grimm V, Berger U, DeAngelis DL, Polhill JG, Giske J, Railsback SF. The ODD protocol: a review and first update. *Ecological modelling*. 2010 Nov 24;221(23):2760-8.

Guisan A, Tingley R, Baumgartner JB, Naujokaitis-Lewis I, Sutcliffe PR, Tulloch AI, Regan TJ, Brotons L, McDonald-Madden E, Mantyka-Pringle C, Martin TG. Predicting species distributions for conservation decisions. *Ecology letters*. 2013 Dec 1;16(12):1424-35.

Gutierrez AP, Ponti L, Dalton DT. Analysis of the invasiveness of spotted wing *Drosophila* (*Drosophila suzukii*) in North America, Europe, and the Mediterranean Basin. *Biological Invasions*. 2016 Dec 1;18(12):3647-63.

Hauser M. A historic account of the invasion of *Drosophila suzukii* (Matsumura)(Diptera: Drosophilidae) in the continental United States, with remarks on their identification. *Pest management science*. 2011 Nov 1;67(11):1352-7.

Hellmann JJ, Byers JE, Bierwagen BG, Dukes JS. Five potential consequences of climate change for invasive species. *Conservation biology*. 2008 Jun 1;22(3):534-43.

Hilborn R, Mangel M. *The ecological detective: confronting models with data*. Princeton University Press; 1997. pp. 145-157.

Hobbs NT, Hilborn R. Alternatives to statistical hypothesis testing in ecology: a guide to self teaching. *Ecological Applications*. 2006 Feb 1;16(1):5-19.

Hoffmann AA, Sørensen JG, Loeschke V. Adaptation of *Drosophila* to temperature extremes: bringing together quantitative and molecular approaches. *Journal of Thermal Biology*. 2003 Apr 30;28(3):175-216.

Hoffmann AA. Acclimation for desiccation resistance in *Drosophila melanogaster* and the association between acclimation responses and genetic variation. *Journal of Insect Physiology*. 1990 Dec 31;36(11):885-91.

Houlihan JE, McKinney ST, Anderson TM, McGill BJ. The priority of prediction in ecological understanding. *Oikos*. 2016 Oct 1.

Jakeman AJ, Letcher RA, Norton JP. Ten iterative steps in development and evaluation of environmental models. *Environmental Modelling & Software*. 2006 May 31;21(5):602-14.

Jakobs R, Gariepy TD, Sinclair BJ. Adult plasticity of cold tolerance in a continental-temperate population of *Drosophila suzukii*. *Journal of insect physiology*. 2015 Aug 31;79:1-9.

Johnson JB, Omland KS. Model selection in ecology and evolution. *Trends in ecology & evolution*. 2004 Feb 29;19(2):101-8.

Johnson IR. Modelling Plant and Crop Growth. Environmental Gene Release: Models, experiments and risk assessment. 1994 Jul 31:11.

Kanzawa T, Studies on *Drosophila suzukii* Mats. *J Plant Prot (Tokyo)* 23:66–70, 127–132, 183–191. Abstract in *Rev Appl Entomol* 24: 315 (1936).

Kearney M, Porter W. Mechanistic niche modelling: combining physiological and spatial data to predict species' ranges. *Ecology letters*. 2009 Apr 1;12(4):334-50.

Kendall BE, Briggs CJ, Murdoch WW, Turchin P, Ellner SP, McCauley E, Nisbet RM, Wood SN. Why do populations cycle? A synthesis of statistical and mechanistic modeling approaches. *Ecology*. 1999 Sep 1;80(6):1789-805.

Kendall BE, Ellner SP, McCauley E, Wood SN, Briggs CJ, Murdoch WW, Turchin P. Population cycles in the pine looper moth: Dynamical tests of mechanistic hypotheses. *Ecological Monographs*. 2005 Feb 1;75(2):259-76.

Kimura MT. Quantitative response to photoperiod during reproductive diapause in the *Drosophila auraria* species-complex. *Journal of Insect Physiology*. 1990 Jan 1;36(3):147-52.

Kimura MT. Cold and heat tolerance of drosophilid flies with reference to their latitudinal distributions. *Oecologia*. 2004 Aug 1;140(3):442-9.

Kingsolver JG, Woods HA, Buckley LB, Potter KA, MacLean HJ, Higgins JK. Complex life cycles and the responses of insects to climate change. *Integrative and Comparative Biology*. 2011 Jun 30;icr015.

Kinjo H, Kunimi Y, Nakai M. Effects of temperature on the reproduction and development of *Drosophila suzukii* (Diptera: Drosophilidae). *Applied entomology and zoology*. 2014 May 1;49(2):297-304.

Langille AB, Arteca EM, Newman JA. (2017) The impacts of climate change on the abundance and distribution of the Spotted Wing Drosophila (*Drosophila suzukii*) in the United States and Canada. *PeerJ* 5:e3192 <https://doi.org/10.7717/peerj.3192>

Langille AB, Arteca EM, Ryan GD, Emiljanowicz LM, Newman JA. North American invasion of Spotted-Wing Drosophila (*Drosophila suzukii*): A mechanistic model of population dynamics. *Ecological Modelling*. 2016 Sep 24;336:70-81.

Lavrinenko A, Kesäniemi J, Watts PC, Serga S, Pascual M, Mestres F, Kozeretska I. First record of the invasive pest *Drosophila suzukii* in Ukraine indicates multiple sources of invasion. *Journal of Pest Science*. (2016):1-9.

Lee JC, Bruck DJ, Curry H, Edwards D, Haviland DR, Van Steenwyk RA, Yorgey BM. The susceptibility of small fruits and cherries to the spotted-wing drosophila, *Drosophila suzukii*. *Pest management science*. 2011 Nov 1;67(11):1358-67.

Lee JC, Dreves AJ, Cave AM, Kawai S, Isaacs R, Miller JC, Van Timmeren S, Bruck DJ. Infestation of wild and ornamental noncrop fruits by *Drosophila suzukii* (Diptera: Drosophilidae). *Annals of the Entomological Society of America*. 2015 Feb 5:sau014.

Legates DR, McCabe GJ. Evaluating the use of “goodness-of-fit” measures in hydrologic and hydroclimatic model validation. *Water resources research*. 1999 Jan 1;35(1):233-41.

Liebold AM, Tobin PC. Population Ecology of Insect Invasions and Their Management. *Annu. Rev. Entomol.*. 2008 Jan 7;53:387-408.

Leibold MA. The niche concept revisited: mechanistic models and community context. *Ecology*. 1995 Jul 1;76(5):1371-82.

Little CM, Chapman TW, Moreau DL, Hillier NK. Susceptibility of Selected Boreal Fruits and Berries to the Invasive Pest *Drosophila suzukii* (Diptera: Drosophilidae). *Pest Management Science*. 2016 Jul 1.

Lumme J, Oikarinen A, Lakovaara S, Alatalo R. The environmental regulation of adult diapause in *Drosophila littoralis*. *Journal of insect physiology*. 1974 Oct 1;20(10):2023-33.

Maurer EP, Brekke L, Pruitt T, Duffy PB. Fine-resolution climate projections enhance regional climate change impact studies. *Eos, Transactions American Geophysical Union*. 2007 Nov 20;88(47):504-.

- Marsh HW, Hau KT. Assessing goodness of fit: Is parsimony always desirable?. *The Journal of Experimental Education*. 1996 Jul 1;64(4):364-90.
- Maurer EP, Brekke L, Pruitt T, Duffy PB. Fine-resolution climate projections enhance regional climate change impact studies. *Eos, Transactions American Geophysical Union*. 2007 Nov 20;88(47):504-.
- Mayer DG, Butler DG. Statistical validation. *Ecological modelling*. 1993 Jul 1;68(1-2):21-32.
- Mayer RE. Models for understanding. *Review of educational research*. 1989 Mar 1;59(1):43-64.
- Mitsui H, Beppu K, Kimura MT. Seasonal life cycles and resource uses of flower-and fruit-feeding drosophilid flies (Diptera: Drosophilidae) in central Japan. *Entomological Science*. 2010 Mar 1;13(1):60-7.
- Morin A, Urban J, Sliz P. A quick guide to software licensing for the scientist-programmer. *PLoS Comput Biol*. 2012 Jul 26;8(7):e1002598.
- Nathan R, Safriel UN, Noy-Meir I. Field validation and sensitivity analysis of a mechanistic model for tree seed dispersal by wind. *Ecology*. 2001 Feb 1;82(2):374-88.
- Newman JA. Using the output from global circulation models to predict changes in the distribution and abundance of cereal aphids in Canada: a mechanistic modeling approach. *Global Change Biology*. 2006 Sep 1;12(9):1634-42.
- Newman JA, Anand M, Henry HA, Hunt SL, Gedalof Z. *Climate change biology*. CABI; 2011.
- Newman JA, Gibson DJ, Parsons AJ, Thornley JH. How predictable are aphid population responses to elevated CO<sub>2</sub>? *Journal of Animal Ecology*. 2003 Jul 1;72(4):556-66.
- Olesen JE, Bindi M. Consequences of climate change for European agricultural productivity, land use and policy. *European journal of agronomy*. 2002 Jun 30;16(4):239-62.
- Ohtsu T, Katagiri C, Kimura MT, Hori SH. Cold adaptations in *Drosophila*. Qualitative changes of triacylglycerols with relation to overwintering. *Journal of Biological Chemistry*. 1993 Jan 25;268(3):1830-4.
- Overgaard J, Hoffmann AA, Kristensen TN. Assessing population and environmental effects on thermal resistance in *Drosophila melanogaster* using ecologically relevant assays. *Journal of Thermal Biology*. 2011 Oct 31;36(7):409-16.
- Parmesan C. Ecological and evolutionary responses to recent climate change. *Annual Review of Ecology, Evolution, and Systematics*. 2006 Jan 1:637-69.
- Parmesan C, Yohe G. A globally coherent fingerprint of climate change impacts across natural systems. *Nature*. 2003 Jan 2;421(6918):37-42.
- Pimentel D, Zuniga R, Morrison D. Update on the environmental and economic costs associated with alien-invasive species in the United States. *Ecological Economics*. 2005;52:273-88.

Pollino CA, Woodberry O, Nicholson A, Korb K, Hart BT. Parameterisation and evaluation of a Bayesian network for use in an ecological risk assessment. *Environmental Modelling & Software*. 2007 Aug 31;22(8):1140-52.

Power M. The predictive validation of ecological and environmental models. *Ecological modelling*. 1993 Jul 31;68(1):33-50.

Prasad AM, Iverson LR, Peters MP, Bossenbroek JM, Matthews SN, Sydnor TD, Schwartz MW. Modeling the invasive emerald ash borer risk of spread using a spatially explicit cellular model. *Landscape ecology*. 2010 Mar 1;25(3):353-69.

Preacher KJ. Quantifying parsimony in structural equation modeling. *Multivariate Behavioral Research*. 2006 Sep 1;41(3):227-59.

Reichert P, Omlin M. On the usefulness of overparameterized ecological models. *Ecological Modelling*. 1997 Feb 14;95(2):289-99.

Ryan GD, Emiljanowicz L, Wilkinson F, Kornya M, Newman JA. Thermal Tolerances of the Spotted-Wing Drosophila Drosophila suzukii (Diptera: Drosophilidae). *Journal of economic entomology*. 2016 Feb 14:tow006.

Salminen TS, Vesala L, Hoikkala A. Photoperiodic regulation of life-history traits before and after eclosion: egg-to-adult development time, juvenile body mass and reproductive diapause in Drosophila montana. *Journal of insect physiology*. 2012 Dec 31;58(12):1541-7.

Saryazdi S, Cheriet M. PKCS: A polynomial kernel family with compact support for scale-space image processing. *IEEE transactions on image processing*. 2007 Sep;16(9):2299-308.

Saunders DS, Henrich VC, Gilbert LI. Induction of diapause in *Drosophila melanogaster*: photoperiodic regulation and the impact of arrhythmic clock mutations on time measurement. *Proceedings of the National Academy of Sciences*. 1989 May 1;86(10):3748-52.

Shearer PW, West JD, Walton VM, Brown PH, Svetec N, Chiu JC. Seasonal cues induce phenotypic plasticity of *Drosophila suzukii* to enhance winter survival. *BMC ecology*. 2016 Mar 22;16(1):1.

Stephens AR, Asplen MK, Hutchison WD, Venette RC. Cold hardiness of winter-acclimated *Drosophila suzukii* (Diptera: Drosophilidae) adults. *Environmental entomology*. 2015 Dec 1;44(6):1619-26.

Thomas MB. Ecological approaches and the development of “truly integrated” pest management. *Proceedings of the National Academy of Sciences*. 1999 May 25;96(11):5944-51.

Storer NP. A spatially explicit model simulating western corn rootworm (Coleoptera: Chrysomelidae) adaptation to insect-resistant maize. *Journal of economic entomology*. 2003 Oct 1;96(5):1530-47.

Teixeira EI, Fischer G, van Velthuizen H, Walter C, Ewert F. Global hot-spots of heat stress on agricultural crops due to climate change. *Agricultural and Forest Meteorology*. 2013 Mar 15;170:206-15.

Thornley JH. Grassland dynamics: an ecosystem simulation model. CAB international; 1998:xi

Thornley JH, Johnson IR. Plant and crop modeling; a mathematical approach to plant and crop physiology. Clarendon. 1990:3-9.

Thornley JH, France J. Mathematical models in agriculture: quantitative methods for the plant, animal and ecological sciences. Cabi; 2007.

Tochen S, Woltz JM, Dalton DT, Lee JC, Wiman NG, Walton VM. Humidity affects populations of *Drosophila suzukii* (Diptera: Drosophilidae) in blueberry. *Journal of Applied Entomology*. 2016 Feb 1;140(1-2):47-57.

Tochen S, Dalton DT, Wiman N, Hamm C, Shearer PW, Walton VM. Temperature-related development and population parameters for *Drosophila suzukii* (Diptera: Drosophilidae) on cherry and blueberry. *Environmental entomology*. 2014 Apr 1;43(2):501-10.

Topping CJ, Høye TT, Olesen CR. Opening the black box—Development, testing and documentation of a mechanistically rich agent-based model. *Ecological Modelling*. 2010 Jan 24;221(2):245-55.

Vandekerckhove J, Matzke D, Wagenaars EJ. Model comparison and the principle of parsimony. 2014

Van Vuuren DP, Edmonds J, Kainuma M, Riahi K, Thomson A, Hibbard K, Hurtt GC, Kram T, Krey V, Lamarque JF, Masui T. The representative concentration pathways: an overview. *Climatic change*. 2011 Nov;109:5-31.

Wallingford AK, Loeb GM. Developmental Acclimation of *Drosophila suzukii* (Diptera: Drosophilidae) and Its Effect on Diapause and Winter Stress Tolerance. *Environmental Entomology*. 2016 Aug 1;45(4):1081-9.

Vinatier F, Tixier P, Le Page C, Duyck PF, Lescourret F. COSMOS, a spatially explicit model to simulate the epidemiology of *Cosmopolites sordidus* in banana fields. *Ecological Modelling*. 2009 Sep 24;220(18):2244-54.

Umina PA, Weeks AR, Kearney MR, McKechnie SW, Hoffmann AA. A rapid shift in a classic cline pattern in *Drosophila* reflecting climate change. *Science*. 2005 Apr 29;308(5722):691-3.

Wallingford AK, Lee JC, Loeb GM. The influence of temperature and photoperiod on the reproductive diapause and cold tolerance of spotted-wing drosophila, *Drosophila suzukii*. *Entomologia Experimentalis et Applicata*. 2016 Jun 1;159(3):327-37.

Wallingford AK, Loeb GM. Developmental Acclimation of *Drosophila suzukii* (Diptera: Drosophilidae) and Its Effect on Diapause and Winter Stress Tolerance. *Environmental Entomology*. 2016 Aug 1;45(4):1081-9.

Walse SS, Krugner R, Tebbets JS. Postharvest treatment of strawberries with methyl bromide to control spotted wing drosophila, *Drosophila suzukii*. *Journal of Asia-Pacific Entomology*. 2012 Sep 30;15(3):451-6.

Walsh DB, Bolda MP, Goodhue RE, Dreves AJ, Lee J, Bruck DJ, Walton VM, O'Neal SD, Zalom FG. *Drosophila suzukii* (Diptera: Drosophilidae): invasive pest of ripening soft fruit expanding its geographic range and damage potential. *Journal of Integrated Pest Management*. 2011 Apr 1;2(1):G1-7.

Ward NL, Masters GJ. Linking climate change and species invasion: an illustration using insect herbivores. *Global Change Biology*. 2007 Aug 1;13(8):1605-15.

Warren DL, Seifert SN. Ecological niche modeling in Maxent: the importance of model complexity and the performance of model selection criteria. *Ecological Applications*. 2011 Mar 1;21(2):335-42.

Way MJ, Van Emden HF. Integrated pest management in practice—pathways towards successful application. *Crop protection*. 2000 Mar 31;19(2):81-103.

Wiman NG, Walton VM, Dalton DT, Anfora G, Burrack HJ, Chiu JC, Daane KM, Grassi A, Miller B, Tochen S, Wang X. Integrating temperature-dependent life table data into a matrix projection model for *Drosophila suzukii* population estimation. *PLoS One*. 2014 Sep 5;9(9):e106909.

Zavalloni C, Andresen JA, Flore JA. Phenological Models of Flower Bud Stages and Fruit Growth of Montmorency Sour Cherry Based on Growing Degree-day Accumulation. *Journal of the American Society for Horticultural Science*. 2006 Sep 1;131(5):601-7.

Ziter C, Robinson EA, Newman JA. Climate change and volitinism in Californian insect pest species: sensitivity to location, scenario and climate model choice. *Global change biology*. 2012 Sep 1;18(9):2771-80.

Ziska LH, Blumenthal DM, Runion GB, Hunt Jr ER, Diaz-Soltero H. Invasive species and climate change: an agronomic perspective. *Climatic Change*. 2011 Mar 1;105(1-2):13-42.

## Appendix 1: Chapter 2 Supplemental Figures and Model Comparison

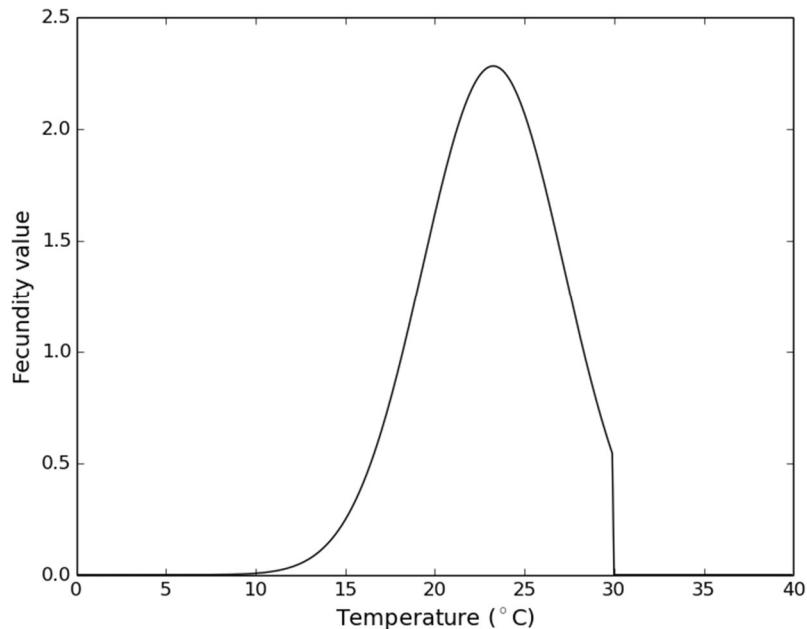


Figure A1.1  $f_T$  (eqn 1) for a range of temperature values between 0 $^{\circ}\text{C}$  and 40 $^{\circ}\text{C}$ .

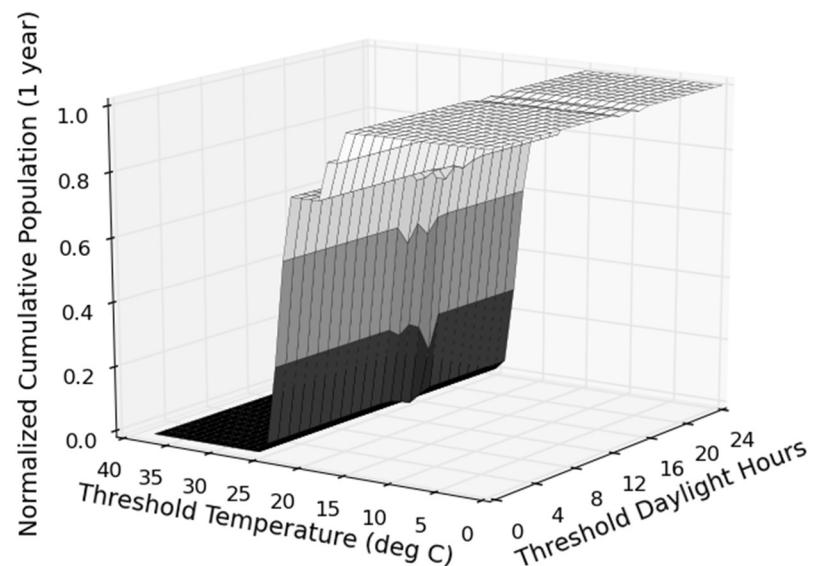


Figure A1.2 The effect of both diapause threshold daylight hours and diapause threshold temperature on the population size.

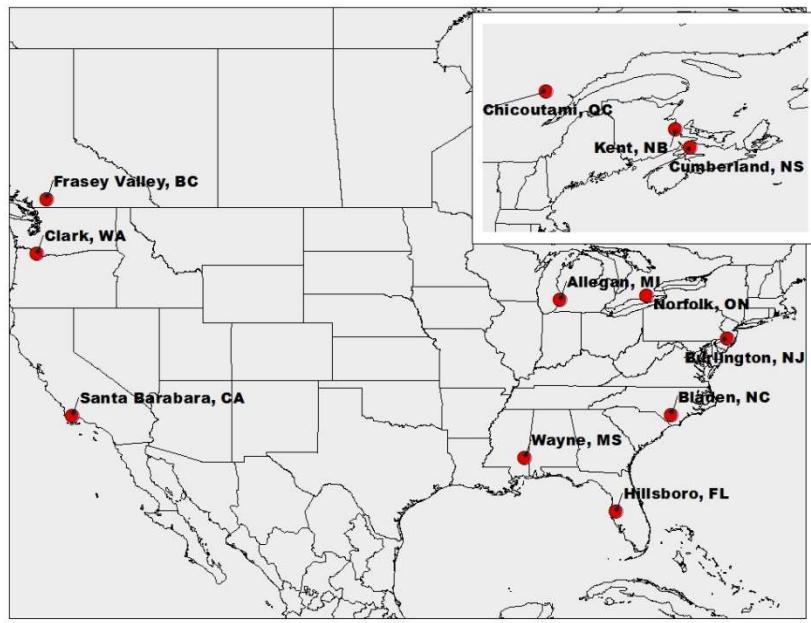


Figure A1.3 A map of the locations in the United States and Canada for which observed temperature data was obtained and used to parameterized the presented model.

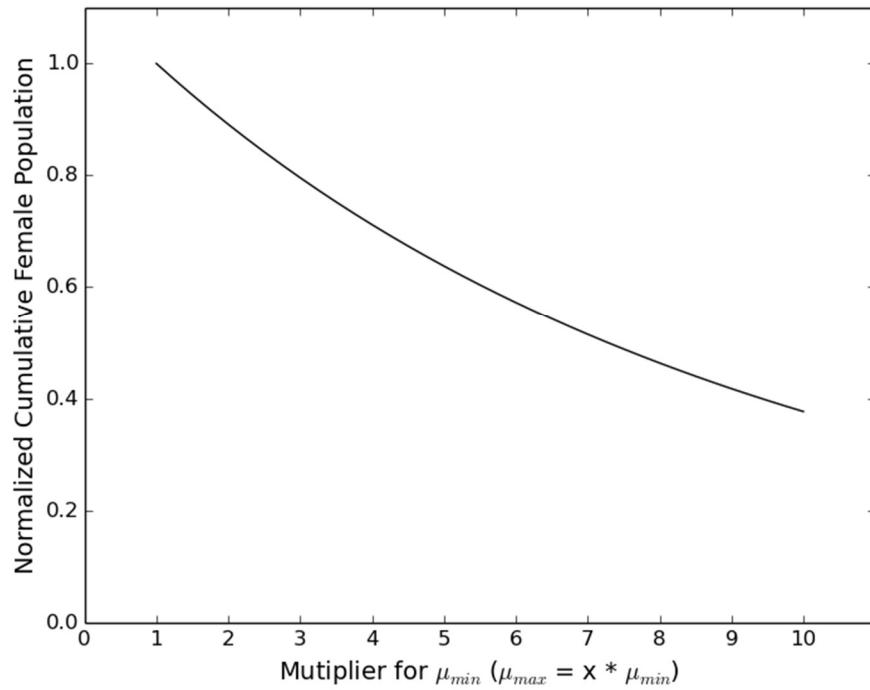


Figure A1.4 Sensitivity tests on the intrinsic mortality parameters when the base model, diapause sub model and fruit sub model are all considered. Increasing the multiplier for  $\mu_{max}$  does increase the resulting population size but does not change the shape of the population growth curve over time.

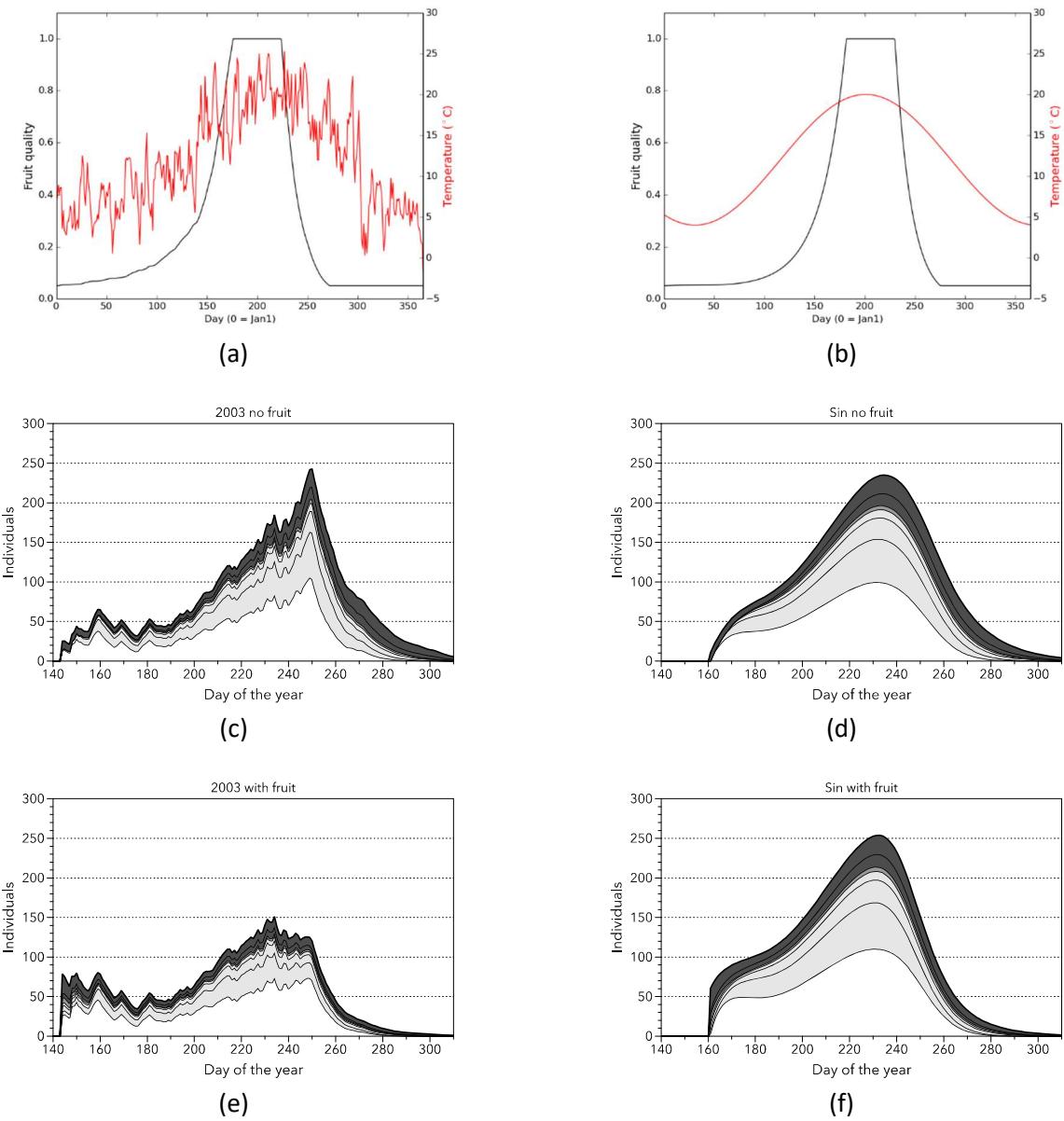
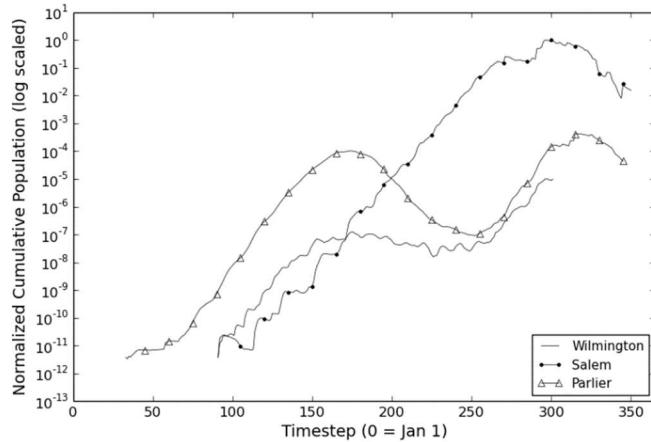


Figure A1.5 (a) temperature profile for Clark WA, 2003 and (b) a simulated temperature curve (see Chapter 2 text, Figure 2.3, profile B). Fruit quality index in (a) and (b) was parameterized using  $gt = 4$ ,  $h = 50$ ,  $\phi = 1$  and  $\lambda = 0.1$ . Stage specific population structure (c through f) for simulations using different temperature profiles (a and b). Population stages from top to bottom are females (all stages combined), males, pupae, instar 3, instar 2, instar 1 and eggs.

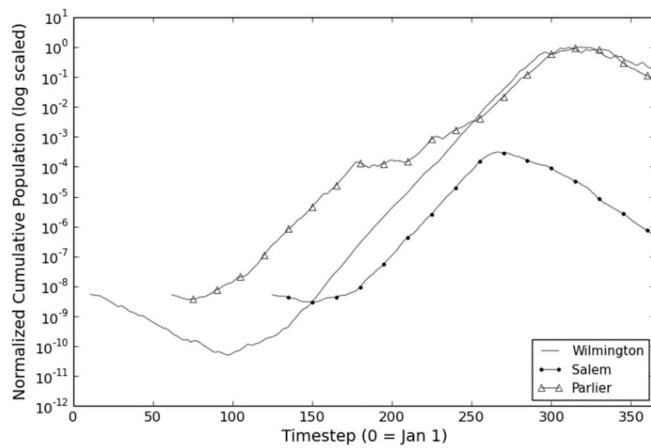
## Model Comparison – Wiman (2014)

For the sake of comparison, I configured our model with the 2013 temperature data for three of the locations discussed in Wiman et al. (Figure 12a); Salem OR, Wilmington NC and Parlier CA. I simulated the population growth for each of these locations using two different configurations. The first configuration (Figure 12b) ignores both the diapause and fruit quality sub models for a direct comparison with the Wiman model while the second configuration (Figure 12c) has both of these sub models enabled.

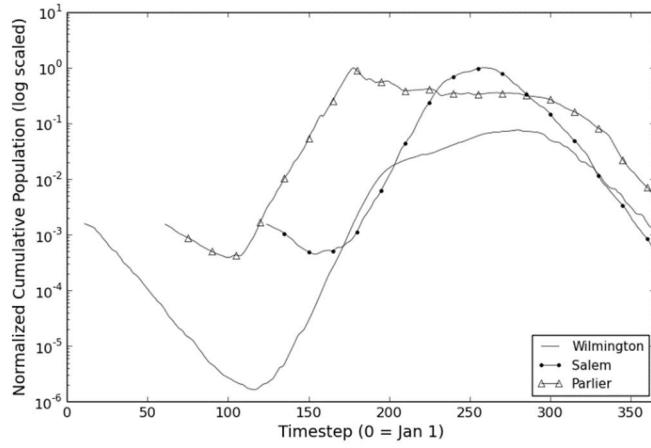
While the results appear to be similar it is worth noting some important differences. In the Wiman model, the population for Salem increases through the year peaking in mid-October and declining for the remainder of the simulation. Our model produces a similar curve but not for Salem, instead this growth is seen for both Wilmington and Parlier. When considering the 2013 temperature profile for the three locations, it is unclear how Salem could sustain such a level of population growth given late autumn temperatures and the observed development rates at these temperatures. (cf. Tochen et al. 2014, Fig 3). Our model inhibits the growth of Salem's populations due to cooler temperatures and sub-optimal development at all life stages. Conversely, our model shows sustained population growth in Wilmington and California over the warmer months perhaps somewhat inhibited by extreme temperature mortality.



(a)



(b)



(c)

Figure A1.6 The model presented herein parameterized for comparison with the model of Wiman et al. (2014). Data for (a) provided by Wiman et al. for replication here. Temperature data is 2013 mean daily for Wilmington NC, Salem OR and Parlier CA. Diapause and fruit quality sub models are ignored (b). The model run for 2013 temperatures with the fruit and diapause sub models included (c).

## Summary Model Parameters and Normalization Maxima

	Maximum for Normalization	Diapause Sub model	Fruit Sub model
Figure 4	1.54E+27	off	off
Figure 5	3.49E+09	off	off
Figure 6	5.15E+18	on	off
Figure 7	1.59E+07	on	off
Figure 8	8.70E+04	on	off
Figure 9	3.24E+05	off	on
Figure 10	7.43E+10	on	off and on (see caption)
Figure 11	2.77E+05	off	off
Figure S6b	1.84E+09	off	off
Figure S6c	6.34E+03	on	on

Table A1.1 Model result maxima used for normalization and sub model inclusion state per figure. All simulations were started with 100 females unless otherwise stated.

Diapause Sub model Parameters	
Critical Temperature ( $T_D$ )	18
Daylight Hours ( $h_D$ )	10

Table A1.2 Diapause sub model parameters constant for all simulations unless otherwise stated.

Fruit Sub model Parameters	
$n_{fd}$	4
$m_{fd}$	0.75
Harvest delay time ( $h_{Fr}$ )	50
Base temperature ( $T_{base}$ )	4
Fruit sub model effect multiplier ( $\omega$ )	4
maximum fruit quality before harvest ( $\phi$ )	0.95
Rate of fruit quality decline after harvest ( $\gamma$ )	0.1

Table A1.3 Fruit sub model parameters constant for all simulations unless otherwise stated.

## Appendix 2: Table of Model Parameters

Eqn #	Parameter Name	Parameter Value	Unit of Measure	Description/Extra Notes	
Eqn 1	$T_{f,max}$	30	°C	Eqn 1 fit using nonlinear regression (GraphPad/Prism software) to polynomial with compact support (Saryazdi and Cheriet 2007) for data provided by Ryan et al. (2016).	
	$\lambda$	52.68			
	$\delta$	5.88			
	$\tau$	23.26			
	$\alpha$	676.0	dimensionless		
	$\gamma$	88.38			
Eqn 2	$A$	0.04	dimensionless	Eqn 2 fit using interpolation (GraphPad/Prism software) to an asymmetric sigmoidal (Richard's 5-point) function using data estimated from Kimura (1990).	
	$K$	99.8			
	$v$	0.813			
	$Q$	$3.23 \times 10^{-16}$			
	$B$	2.87			
Eqn 3	$T_D$	18	°C	Temperature required to terminate diapause.	
	$h_D$	10	hours <sup>-day</sup>	Number of daylight hours to initiate diapause.	
Eqn 6	$a$ (all stages)	0.0001113	dimensionless	Eqn 6 fit using nonlinear regression (GraphPad Prism/software) to Briere et al. (1999) development function using data provided by Ryan et al. (2016).	
	$T_L$ (all stages)	8.0139	°C		
	$T_U$ (all stages)	30.99			
	$m_1$ (egg)	0.104	days <sup>-1</sup>		
	$m_2$ (instar 1)	0.082			
	$m_3$ (instar 2)	0.112			
	$m_4$ (instar 3)	0.231			
	$m_5$ (pupa)	0.470			
	$T_{min,\mu}$ (all stages)	3	°C		
	$T_{max,\mu}$ (all stages)	33			
Eqn 7	$\beta_{i,1}$ (all stages)	-0.008			
	$\beta_{i,2}$ (all stages)	0.00032			
	$\beta_{i,3}$ (all stages)	-0.000002			
	$\tau$ (all stages)	8.1776			
	$\beta_{1,0}$ (egg)	0.160			
	$\beta_{2,0}$ (instar 1)	0.140			
	$\beta_{3,0}$ (instar 2)	0.085			
	$\beta_{4,0}$ (instar 3)	0.086			
	$\beta_{5,0}$ (pupa)	0.061			
	$\beta_{6,0}$ (male)	0.134			
	$\beta_{7,0}$ (female 1)	0.069			
	$\beta_{8,0}$ (female 2)	0.090			
	$\beta_{9,0}$ (female 3)	0.201			
	$\beta_{10,0}$ (female 4)	0.051			
	$\beta_{11,0}$ (female 5)	0.301			
	$\beta_{12,0}$ (female 6)	0.251			
	$\beta_{13,0}$ (female 7)	0.330			

Table A2.1 Model parameters including values, units of measure and supplemental notes on estimation methods or purpose.

Eqn #	Parameter Name	Parameter Value	Unit of Measure	Description/Extra Notes
Eqn 7 (ctd.)	$\mu_{1,max}$ (egg)	0.328	individual-day	Eqn 7 fit using nonlinear regression (computational solution) to data provided by Emiljanowicz et al. (2014). In the absence of an empirical estimate for $\mu_{i,max}$ values are assumed to be three times their corresponding values estimated at constant optimal temperature.
	$\mu_{2,max}$ (instar 1)	0.269		
	$\mu_{3,max}$ (instar 2)	0.102		
	$\mu_{4,max}$ (instar 3)	0.107		
	$\mu_{5,max}$ (pupa)	0.30		
	$\mu_{6,max}$ (male)	0.097		
	$\mu_{7,max}$ (female 1)	0.054		
	$\mu_{8,max}$ (female 2)	0.012		
	$\mu_{9,max}$ (female 3)	0.450		
	$\mu_{10,max}$ (female 4)	0.0		
	$\mu_{11,max}$ (female 5)	0.750		
	$\mu_{12,max}$ (female 6)	0.600		
	$\mu_{13,max}$ (female 7)	0.837		
Eqn 8	$\omega$	4	dimensionless	Determines the rate of fruit quality increase.
Eqn 9	$\Phi$	0.95	dimensionless	Maximum fruit quality value.
	$\Upsilon$	0.1		Determines the rate of fruit quality decline.
	$h_{Fr}$	50		Number of days for harvest lag
Eqn 10	$T_{base}$	4	°C	Eqn 10 was computationally developed to estimate the degree day model of sour cherries presented in Zavalloni et al. (2006).
Eqn 11	$m_{fd}$	0.75	dimensionless	Eqn 11 Fruit quality effect on development parameters based on the work of Newman et al. (2003).
	$Q_{p,fd,h}$	0.5		
	$n_{fd}$	4.0		
Eqn 13	$m_{f\mu}$	0.1	dimensionless	Eqn 13 Fruit quality effect on mortality parameters based on the work of Newman et al. (2003).
	$Q_{p,f\mu}$	0.5		
	$n_{f\mu}$	4.0		
Eqn 15	$v_1$	0.832		Parameter data obtained from Emiljanowicz et al. (2014).
	$v_2$	0.807		
	$v_3$	0.763		
	$v_4$	0.556		
	$v_5$	0.324		
	$v_6$	0.257		
	$v_7$	0		
Eqn 18 & Eqn 19	$\theta$	0.5	individual-day	Eqn 18 & 19, $\theta$ is the estimated proportion of flies that are male.
	$d_{F1}$	0.0125		Eqn 19, $d_{Fi}$ is estimated from Emiljanowicz et al. (2014) as the per capita rate of development between female stages representing decline in egg viability.
	$d_{F2}$	0.1		
	$d_{F3}$	0.1		
	$d_{F4}$	0.2		
	$d_{F5}$	0.25		
	$d_{F6}$	0.2		

## Appendix 3: Individual GCM-RCP Population Maps

Individual population “maps” for each of the Global Circulation Model and Relative Concentration Pathway combinations simulated in Chapter 3. Population data are globally normalized across all maps and represent 1 degree latitude by 1 degree longitude.

