Refining a forecasting method for first emergence of an important forestry pest *(Hylobius abietis)* in Ireland through environmental modelling

Cathal Flood B.Sc. (19253490)

Dissertation submitted to the Department of Geography, National University of Ireland, Maynooth University, in partial fulfilment of the requirement for the degree of master's by research (M.Litt.).

Supervisor(s): Prof. Rowan Fealy,

Prof. Christine Griffin and Dr Rafael de Andrade Moral

Submission Date: October 31st, 2023.

Declaration:

I have read and understood the Departmental policy on plagiarism.

I declare that this thesis is my own work and has not been submitted in any form for another degree or diploma at any university or other institution of tertiary education.

Information derived from the published or unpublished work of others included in the research has been acknowledged in the text and a list of references is given.

Signature: C.F

Date: 02/05/2024

Acknowledgements:

“*Think where man's glory most begins and ends, and say my glory was I had such friends*.”
Willam Butler Yeats (1865-1939)

First, I would like to acknowledge funding for this research from the Department of Agriculture, Fisheries, and the Marine as part of the overall project “Towards Integrated pest management for Pine Weevil in Ireland” (DAFM project 17/C/228). I equally should acknowledge the HEA for their covid fund I availed of as well as the understanding and support of Maynooth Universities administration departments. There were many people who offered guidance, assistance, and a helping hand to me completing this project that I must express my sincere gratitude to.

The researchers in this research area in Maynooth before me I must thank for their existing work, calls, and assistance at points of the research including Ben Malone, Aoife Dillon, Louise McNamara, Chris Williams, and Tolis Kapranas. It was much appreciated to discuss their important work that I learned immensely from in Maynooth. I mention too the brilliant people at Forest Research UK including Daegan Inward and Dave Wainhouse who offered me great advice on the work from his forecast model that this research proposed to adapt and grow from. I want to mention Colm Lyons from Coillte for his passionate interest in the work, assistance in showing me existing management protocols and an onsite visit to Kerry pine weevil trial sites in 2021. His enthusiasm in this area is amazing. My Maynooth family of fellow post grads and researchers I want to thank also including Alan Inglis (for his guidance when statistics did not come easy for me), Shirley Chance Howe (for her wise advice on the research life through covid), Gail Maher (for her guidance to not give up) and Azucena Jimenez Casteneda (for her expert help on GIS queries and digital elevation models used in the thesis). Maynooth Security are due a call out for the many hours they let me utilise the access to Laraghbryan House to complete my work. My fellow students and friends in Maynooth Biology, Maths & Stats and Geography deserve mention too for been a big part of my time there. I want to acknowledge my family and friends who continued to support me, in particular my friend John who is the loveliest of human beings and my friend Ami who passed earlier this year. She was a beacon of light in the dark to many and I hope she is at peace now.

I want to thank my mother for her always well wishes at heart in difficult times and my father too. The lessons my parents thought made me who I am today. My father grew up in a time and situation when third level education was but a pipe dream. I never forget the sacrifices he made for the privilege for me to submit this work today. I input my middle name on the thesis cover in honour to him. To my supervisors who kept believing and telling me to dig in, I thank you beyond words for your continued support through this long-lasting work. Rowan Fealy, for his spot-on advice and knowledge in the area, Christine for her incredible commitment to the work and assistance of help always and to Rafael Moral for being so determined to teach me the basics of R and mechanics of forecast modelling as well as been a constant pillar of friendship throughout my life during this time. Finally, I acknowledge Alessandra Lemos Dos Santos. Her zealous for research in Maynooth is just something to behold and she is an even more kind human being. The hours she gives to her work, her friends and somehow to give time to the modelling in this work in terms of `pineR` was amazing. I am proud to call her a friend. Without my supervisors of course and Alessandra this project would not have been finished. I dedicate the thesis to all the people that got me there. I will never forget that.

Abstract

Large pine weevil (*Hylobius abietis*) is a serious pest of coniferous plantations throughout Northern Europe including in Ireland, causing significant mortality in replanted trees. Replanted trees are new transplants set on clear fell sites. This mortality results in severe economic losses to foresters (Langstrom and Day, 2004). Development of weevils takes place in the stumps of felled conifer trees and emerging adults feed directly on new transplants on site. Currently, young trees are mainly protected by chemical insecticides applied when weevil attack is anticipated. In Integrated Pest Management (IPM), adequate timing of management actions can help reduce the necessity for pesticides, or the amount used. In the case of pine weevil, forecasting the extent of weevil infestation is centred around the process of stump hacking to estimate numbers of weevils developing there (Teagasc, 2020). The research outlined here adapts an existing UK simulation model developed to determine geographic variation in voltinism of pine weevil under climate change (Wainhouse *et al.*, 2014) to forecast timing of first year emergence of pine weevil in Ireland. The model utilises historical temperature data, derived either from nearest synoptic stations (weighted for distance) or interpolated gridded (Walsh, 2012) to forecast cumulative weevil emergence for specific sites and years for which existing biological data of emerging weevil populations were available. Observed and model simulated emergence patterns were compared, both for an early version of the model and a corrected version (adapted with the machine learning algorithm random forest). Site-specific co-variates that affect the model forecast simulations of weevil emergence were identified. Previous research at Maynooth University (Williams et al., unpublished) had demonstrated that the original implementation of the UK model could be used to predict site-specific patterns of weevil emergence based on data from local weather stations. This project builds on these findings, resulting in the development of the `pineR` model, incorporating data from additional sites to detect potential site-specific factors of influence, such as elevation, that could be considered to provide more accurate predictions of first year weevil emergence in Ireland. It also lays the groundwork for future work that would potentially incorporate information on weevil population structure in stumps to create an accessible version of the model using the in-stump values and local weather data to forecast timing of weevil emergence from a specific stage. It is envisaged that `pineR` will ultimately help refine the stump assessment protocol (i.e., stump hacking) currently used by foresters in Ireland.

Abbreviations:

CAP: Common Agricultural Policy.

COFORD: Council for Forest Research and Development.

DAFM: Department of Agriculture, Fisheries, and the Marine.

DEM: Digital Elevation Model.

DHPLG: Department of housing, planning and local government.

FAO: Food and Agriculture Organisation of the United Nations.

GIS: Geographic Information Systems.

IPM: Integrated pest management

LOOCV: Leave one out cross validation.

ML: Machine Learning.

RCD: Root collar diameter.

RPO: Research performing organisations.

RF: Random Forest.

Table of Contents

Chapter 1: Introduction

Insect pests are a major constraint on productivity of agricultural and forestry systems and their impacts are likely to increase as the global climate warms (Lehman *et al.*., 2020). The consequences of pest outbreaks are sizeable with pest species (primarily insects) estimated to be responsible for losses of almost 18% of global crop production (Oerke, 2006). Insect pests have been reported to reduce agricultural yields by as much as 16% before harvest, and to consume a similar amount following harvest (Bebber *et al.*, 2013). In the case of forest pests, they are equally damaging. One pest, the gypsy moth *(Lymantria dispar)*, an introduced forest insect pest in North America, has annual control expenditures exceeding US$ 35 million since 1980 (Wallner, 2000). The yield losses from different categories of pests have been estimated to be US$ 500 billion worldwide (Oerke, *et al.*, 1994; Liebhold and Tobin, 2008). Several insect pests have been shown to defoliate trees and degrade plant biodiversity, threaten commercial forestry, and hinder climate change mitigation through increasing tree stress or mortality leading to increased greenhouse gas emissions (Bradshaw *et al.*, 2016). In terms of the land mass of European forests, the area covered is more than 2 million km^2, accounting for 33% of the continent's land surface. Even though forests are highly resistant ecosystems, they are vulnerable to sudden or extreme changes in environmental conditions as the long lifespan of trees limits their ability to adapt, particularly for a pest outbreak (Forzieri *et al.*, 2021).

1.1 Background: An important forest pest

Forest pests are wide ranging in terms of distribution. In a thematic review looking at forest pests in 25 countries commissioned by the Food and Agriculture Organisation (FAO) of the United Stations, over 54% of pests were recorded in planted forests, 31% in naturally regenerated forests and almost 15% in both forest types (FAO, 2009). Forest pests can also severely affect the environments they inhabit. Pests such as the gypsy moth (*Lymantria dispar)* and mountain pine beetle *(Dendroctonus ponderosae)* have been shown to cause the displacement of native trees, widespread defoliation and mortality disrupting habitat functioning, all of which negatively impact on biodiversity (Fajvan and Wood, 1996; Janes *et al.*, 2014). 86% of pest species identified in the 2009 FAO review of global forest pests and diseases were insect pests with the majority recorded in naturally regenerated forests (almost 62%) (FAO, 2009). One of the most prominent and damaging pests of reforestation and establishment forestry in northern and central Europe is the large pine weevil *(Hylobius abietis)* (Leather *et al.*, 1999; Villamor *et al.*, 2019) which is the focus of this current research.

As a pest of young coniferous and deciduous trees, the large pine weevil is the only forest pest for which prophylactic treatment has become routine practice in the establishment of clear-fell sites. This has led to potentially harmful chemical control utilised in terms of environmental and worker safety as well as efficacy (Stoakley and Heritage, 1990; Dillon *et al.*, 2006). However, it has been estimated that, in the absence of such insecticidal prevention, economic losses due to this pest would be in the region of €140 million across Europe, annually (Langstrom and Day, 2004; Lalik *et al.*, 2021.

1.2 Integrated pest management (IPM) and current strategies (Large Pine Weevil)

Environmental monitoring of pests and diseases has a key role to play in reducing both the economic and environmental impacts of pests and is part of a comprehensive system of Integrated Pest Management (IPM). This includes monitoring and forecasting with the aim of ultimately managing pests in a more environmentally friendly fashion. Monitoring involves regular surveillance of key insect traits from their population, development stages, biology, and movement or migration, amongst other criteria. In certain cases, it may be difficult to ascertain what pests are causing damage to a crop. In this case, symptoms of crop damage must be investigated (Dent, 1995). Pests have been monitored in a variety of ways with various tools developed, such as traps to catch and record insect activity (Prasad and Prabhakar, 2012). Forecasting of pest outbreaks is also essential to IPM, particularly the accurate prediction of severity of pest infestation. In considering how IPM might be applied in any situation, it is important to consider both strategic and tactical approaches to managing a pest situation and that IPM is a whole of the landscape, year-round, approach. A component of this involves the use of biological pesticides or selective synthetic pesticides that do not disrupt species other than the target pest (Fitt and Wilson, 2012). Many pests have traditional treatment applications. Control of the grain aphid (*Sitobion avenae*) for example and its associated viruses has been achieved via the routine application of chemical controls, irrespective of the level of risk from season to season (Duffy *et al.*, 2017). This approach has resulted in both negative ecological effects, as well as the emergence of highly resistant genes in some aphid species considered highly problematic (Foster *et al.*, 2007). The change of European legislation to further enforce pesticide regulations has only further enhanced the importance of forecasting tools to complement current IPM strategies and reduce the use of harmful chemical intervention.

Pest characteristics can also be included in IPM protocols. For instance, in the case of the large pine weevil, the pest is attracted to volatiles or the smell of fresh timber from felled or damaged coniferous trees (Dillon *et al.*, 2006; Nordlander *et al.*, 1986) (see Section 1.3 below). Research in Ireland has shown that with up to 100,000 adult weevils emerging per hectare on recently felled conifer sites, it would not be uncommon to observe 100% mortality of transplant seedlings without the aid of chemical control (Dillon and Griffin, 2008; Dillon *et al.*, 2008). When chemical intervention is utilised, the protection is time limited. For example, the application of cypermethrin to new trees is only effective for up to 6 weeks following application. Current forecasting of weevil emergence in Ireland is centred on the process of stump hacking, counting the number of weevil larvae and pupae in stumps - and applying a threshold to estimate pest emergence in accordance with the time the site was felled (Teagasc, 2017). At an operational level, weather conditions and site elevation are considerations in deciding to spray or not. Given its economic implications, control of pine weevil is essential for sustainable forest operations with the use of pesticides. However, chemical intervention is not completely effective as adult weevils tend to avoid seedlings treated with certain insecticides and seek out untreated seedlings. While feeding can be depressed, death may take several weeks after pesticide poisoning, during which time weevils continue to damage seedlings (Rose *et al*., 2005). Therefore, an improved forecasting method to identify the timing of emergence of adult weevils, especially in the first year when populations are highest, would significantly aid the management process providing more effective decision support tools for more appropriate pest management strategies. Knowing when plants or crops are at risk would facilitate more accurate timing of targeted chemical intervention if needed and minimise the use of such chemicals with potentially harmful consequences to the environment.

Effective decision support tools are required to provide onsite practitioners with advice regarding appropriate pest management strategies for their enterprise (Duffy *et al.*, 2017). Coillte, the commercial forestry body in Ireland owned by the State, has adopted many IPM methods for the large pine weevil (*Hylobius abietis*). Stump hacking, as mentioned above, is used to predict the severity of a weevil outbreak and inform the forester if chemical intervention is required or not. Non-chemical strategies have been adopted including earlier planting, use of more vigorous or thicker planting stock, mounding and feeding barriers around plants (Teagasc, 2020). If insecticide is used, dipped plants are followed up by top-up sprays if required based on the stump hacking results and monitoring of treated and untreated plots (Lyons, Coillte, *Pers. Comm.,* 2021). Extensive research has also been undertaken on the potential for using -biological control agents in pine weevil management in Ireland looking at application methods, effects of soil type on efficacy, and operational use of entomopathogenic nematodes (EPNs) for instance (e.g., Williams *et al*., 2013; Dillon *et al.,* 2006; Kapranas *et al.,* 2017). Significant research has also been undertaken in the UK that has produced many useful findings for pine weevil management (Moore, 2004; Inward *et al.*, 2012. Wainhouse *et al.*, 2014). More recently an overview document on guidance for IPM of *Hylobius abietis* in UK forestry was completed (Willoughby *et al*., 2017). This report has relevance for Ireland but there are differences in silvicultural methods between the UK and Ireland which must be noted (Fedderwitz *et al*. 2023). Firstly, there is less regional variation in the life cycle duration in Ireland with less likelihood of univoltine cycles compared to the UK. Secondly, nurseries in the UK provide treatment via insecticides by spraying the seedlings whereas general practice in Irish nurseries is dipping. Thirdly, a fallow period of at least five years is outlined in the strategy or the UK, but in Ireland, felling licenses strictly limit the fallow time. Finally, due to the difference in fallow periods and the use of billet trap counts of weevil populations, the *Hylobius* decision system largely used in the UK to predict weevil infestation on a clear cut is not directly applicable in Ireland (Fedderwitz *et al*. 2023).

1.3 Life cycle and phenological modelling of the large pine weevil (*Hylobius abietis*)

Adult weevils first arrive via migration to clear-fell sites, attracted by the aroma of freshly cut wood (Dillon *et al*., 2008). The females oviposit in the stumps of recently felled coniferous trees, where immature weevils start to develop beneath the bark. Eggs hatch into larvae that appear a creamy white colour with a brown head capsule. Larvae feed under the bark of tree roots and stump, where they pass through four larval moults before pupation (Nordenham and Nordlander, 1994). In colder climates, larval development habitually can take up to two years and can be as long as five years. Typically, 75% of larvae develop within one year (Bejer-Peterson *et al.*, 1962). The larvae develop into pupae which are immobile, creamy white in colour and soft bodied. The stumps of most coniferous trees can support development, but the populations developing in the tree stumps and emerging from pine (i.e., *Pinus* species) are much higher than from spruce (i.e., *Picea* species) (Dillon and Griffin, 2008). The size of populations colonising clearfell sites from adjacent standing forests is largely dictated by the availability of food in these standing forests or access to fresh stumps. The practice of clear felling provides adult female weevils with breeding sites to oviposit in fresh coniferous stumps. Clear felling, where all trees are felled at once, is the standard forestry practice in Ireland. In non-clearfell areas, only some trees are felled

- at any one time (Inward *et al*., 2012). Figure 1.1 shows the life cycle of pine weevil infestation on a clear fell site.

Figure 1.1: The large pine weevil life cycle on a clear fell site from initial migration and oviposition by females on freshly felled tree stumps to development and subsequent generations whilst feeding on new transplants, both coniferous and deciduous.

Adult weevils emerge from pupal cells at the end of summer (July to September), although a large proportion overwinter within the pupal chamber also (Leather *et al*., 1999). Weevils develop from egg to adult in 12-36 months in Ireland and the U.K, with development typically taking longer in spruce compared to pine and with notably slower growth rates. The late summer to early autumn is the peak period of emergence for pine weevil (Dillon and Griffin, 2008; Teagasc, 2020). Emerging adults can remain on site for a few weeks but have been found to migrate to nearby sites that have been recently felled. Following emergence, adult weevils feed extensively on tree seedlings or transplants (both coniferous and deciduous) on site resulting in significant damage (Leather et al., 1999; Wainhouse *et al.*, 2007). In response, the current protocol for chemical treatment recommends intervention for regions within a 5 km radius of a neighbouring clear-fell site (Dillon *et al*., 2008; Teagasc, 2020). Weevils are normally active between March and October and can live for up to four years, going to ground to hibernate in the soil litter interface of mature forest stands during the winter (Munro, 1928; Leather *et al*., 1999). Pine weevils are semivoltine meaning they produce generations or broods less than once per -

year. Their life cycle is an important determinant of economic impact as a pest influencing the length of a fallow period or the frequency of insecticide use imposed on a managed clear fell site.

In Ireland and the UK, the large pine weevil typically has a 2-year life cycle, but generation time (voltinism) across its broad Palaearctic range can vary from a single year (univoltine) in the warmer southern forests, to up to 4 years in the cooler north. This indicates a pliable life cycle which is reactive to local conditions (Wainhouse *et al.*, 2007). Increasing temperatures, due to climate change, are likely to affect many insects and lead to complex responses of insect phenology with noticeable shifts towards earlier seasonal activity. Voltinism can increase with warming, and diapause schedules may also be impacted due to greater thermal accumulation before the critical photoperiod (daylength) is reached. Ultimately, forecasting changes in phenology requires improved understanding of insect life histories (Forest, 2016). The duration of the weevil life cycle is likely to be affected by such changes, resulting in alterations in the seasonal timing of adult activity, which will in turn affect the economic impact of the pest and the way it is controlled. Forest Research UK undertook research to determine the effect of temperature on pine weevil growth and development (Inward *et al.*, 2012). Their research found that the development rate of the eggs, larvae, and pupae increased linearly with temperature but that development of prepupal larvae was highly variable, as that stage tends to undergo a facultative diapause, which was initiated by development temperature. Weevils reared at higher temperatures grow faster and are larger as shown in Figure 1.2.

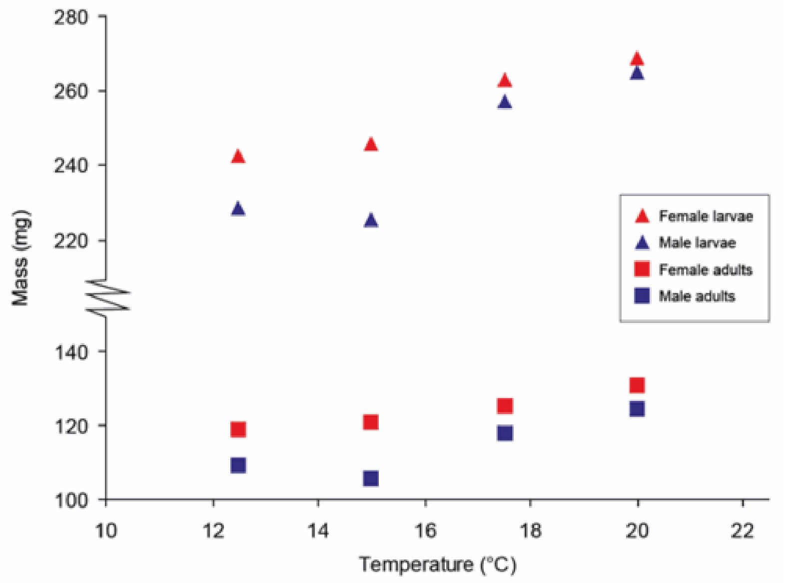


Figure 1.2 The effect of temperature and sex on mean pine weevil mass. Final mass of larvae (top) and adults below is shown (Inward *et al.*, 2012).

Wainhouse *et al.* (2014) developed a model for the UK to simulate the development of weevils through their life cycle, to predict voltinism based on temperature in the root-stump microhabitat. The model was developed to predict changes in voltinism due to projected future changes in the climate system. Modelled future temperatures were obtained from the UK Climate Projections 2009 generator (UKCP09) for a period from the 2030s to the 2070s. The study showed that a warming climate will extend the periods during which weevils can be damaging (Wainhouse *et al*., 2014), further emphasising the need for a refined forecasting method for this forest pest.

As part of a recent INTERREG project (IMPACT), the simulation model developed in the UK was adapted to estimate site-specific patterns of weevil emergence for a selection of sites in Ireland based on observed temperature data from local weather stations. Using empirical data on timing of weevil emergence (assessed by emergence traps erected on clear-fell sites) it was found that in half of the trials, predictions did not differ significantly from the observational data (Williams *et al*., unpublished). The present research seeks to build on that work by examining additional factors, including site-specific factors, to see if the model can be enhanced and ultimately aid the development of a more accurate forecast model for use in decision making under current weather conditions.

Insects as ectothermic organisms are among the group of organisms most likely to be affected by weather and climate with temperature noted as the dominant abiotic factor with a direct influence on their life cycle development (Bale *et al*., 2002). The results from pest simulation models can be most effectively used whilst following correct biological, environmental, experimental, and economic inputs to analyse the most effective management options (Strand, 2000). This thesis aims to build on existing research to refine the forecasting method for the emergence of pine weevil in Ireland. This was ultimately achieved through the development of a weevil forecasting package developed for the R programming language called `pineR` (Lemos dos Santos *et al.* 2023) and will be outlined in detail (See chapter 3).

1.4 Research focus and objectives

This thesis focuses on developing an existing simulation model incorporating additional data and information from a substantial body of research on pine weevil emergence in Irish forestry, collected during field trials of biocontrol agents conducted by Maynooth University and Coillte. This project builds on previous research in Maynooth (Williams et al., unpublished) to investigate how site-specific covariates, such as elevation and soil type, can be included in the model via appropriate statistical tools to identify factors of influence. Observed temperature data from meteorological stations ("met stations") and gridded climate data inputs were employed in the simulation model initially to forecast first year emergence and compare this with observed biological emergence data. Model simulated cumulative emergence of weevils were then compared with the available empirical data using appropriate statistical techniques to cross-validate the model. Geographical site-specific conditions were examined also to see can they help provide more accurate predictions of emergence times. Population structure data from stump hacking is also used to compare between sites. This research ultimately examined the potential to refine a life cycle simulation model by taking account of important site covariate factors to forecast emergence of pine weevil. The final iteration of the forecast model implemented, `pineR,` was achieved with the following objectives.

- Employ phenological data from 27 sites (emergence trial site locations) as model input along with local temperature data for the years of the trial, sourced from the nearest meteorological stations and weighted for distance, to compare observed and simulated emergence in a life cycle model.
- Employ obtained gridded climate data (0.01-degree resolution; (Met Eireann 1km grids; Walsh, 2012)), derived from the existing network of meteorological observations using spatial interpolation techniques, to drive a finalised phenological forecast model for Ireland.
- Examine site-specific geographical data as covariates in the model to see do factors such as elevation, soil type and slope explain any mismatches between the observed and the modelled data. Factors that influence the accuracy of the model in forecasting time of emergence can be considered in a corrected version of the model to improve accuracy of forecast.

- Evaluate supplementary data on weevil population structures in stumps including the proportion of weevils at each developmental stage (i.e., larvae, pupae, adult) for sites at which emergence was modelled. Population size of weevils on trial sites will also be examined for comparison.
- Refine the existing simulation model using emergence data and temperature data inputs for Irish trial sites with a proposed correction to forecast timing of weevil emergence to inform pest protection protocols for pine weevil.

All the project objectives will combine to assess whether the accurate forecasting of weevil first emergence times can be achieved in support of more refined decision making for management of the pest in Ireland.

1.5 Context and value of this research

Given the significant and damaging impact that the large pine weevil causes the forest sector in Ireland and beyond, the need for an appropriate and improved future IPM strategy is clear (Dillon *et al.*, 2006; Wainhouse *et al.*, 2014; Langstrom and Day, 2004). This research aims to create an improved forecasting tool via an emergence forecast model to aid foresters and inform key management decisions such as timings of plant protection product applications based on improved predictions of emergence and development patterns. Fallow periods have been identified as a potential viable future strategy for pine weevil management in certain locations but with sites having to be replanted within two years of felling in Ireland, greenplanting (planting large stock immediately after felling) is a more realistic and trending development in Irish forestry. Forests play an important economic, environmental, and recreational role in Ireland today making a significant impact contribution to the Irish economy, currently valued at €2.3 billion, and an increasingly important role in rural development contributing to rural stabilisation and viability. Forestry also has a key role in climate change mitigation as part of the national climate strategy (DAFM, 2023). The research forms part of an overall project on integrated pest management of pine weevil in Ireland that will develop a critical mass of IPM for the forest pest in Ireland.

1.6 Structure of thesis

The thesis is divided into five chapters, each addressing an aspect of the research project undertaken.

Chapter one introduces the topic of pest management including a background on pests, the life cycle of pine weevil, the use of phenological models and current IPM strategies. The chapter also establishes the research focus and objectives, the context value of the research and the structure for the research thesis.

Chapter two reviews existing literature of pest forecasting with an emphasis on pest monitoring and modelling. The history of such methods incorporated into IPM principles will be examined in an introduction to pest forecasting. The relationship between temperature and development/phenology looking at the direct influence of how temperature impacts the development, reproduction, distribution, and survival of insects will be discussed in detail. The development and application of pest forecast models will be outlined looking at benefits and limitations of different methods. The influence of regional specific factors will be reviewed in terms of potential variable importance on pest modelling. Finally, the influence of climate change on Insect development will be examined prior to a conclusion of the chapter.

Chapter three describes the various data types and methodologies used to develop the forecast simulation tool. These include the 27 forest trial sites where pine weevil phenological data (emergence trap and hacking data) was recorded and the temperature data, both the local synoptic weather stations from which Met Eireann historical temperature data were sourced for specific year of trial, and temperatures derived from the Met Eireann grid. The use of the model and any subsequent post processing steps (such as weighted averaging of met station data and cross validation methods) are explained.

Chapter four presents the results. The chapter begins with a look at the observed emergence data on trial sites used in modelling. It then evaluates the met-station temperature data against the gridded climate data eventually used to drive the forecasting model. The model results will be shown with and without a machine-learning correction using both types of temperature inputs. Selected sites of interest will be examined in more detail to help explain mismatches between observed and model-simulated weevil emergence. A section on “variable importance” will investigate the influence of site-specific covariates in the forecast model as identified by machine learning. Additional data e.g., population structure of weevils in stumps will be covered where relevant when discussing the model performance on trial sites. The proposed corrected finalised model results will conclude the chapter.

Chapter five concludes the thesis, discussing the core aspects of the results obtained and findings deduced as well as providing recommendations for any future research or policy application from this refined forecasting method to aid IPM of pine weevil in Ireland in general.

Chapter 2: Literature Review

"*If all mankind were to disappear, the world would regenerate back to the rich state of equilibrium that existed ten thousand years ago. If insects were to vanish, the environment would collapse into chaos.*"

Edward O. Wilson, American Biologist, naturalist, and writer.

This chapter examines existing knowledge on pest forecasting, detailing pest monitoring and modelling. The history of such methods and how they link to Integrated Pest Management (IPM) principles will be outlined. This will be followed by a discussion on the relationship between temperature and development looking at the direct influence of how temperature impacts the development, reproduction, distribution, and survival of insects. The development and application of pest forecast models will be reviewed looking at benefits and limitations of current practices and how IPM of insect pests can be refined for use in forecasting methods. The influence of regional specific factors will be reviewed in terms of variable importance on pest modelling with an emphasis on the large pine weevil *(Hylobius abietis)*. Finally, the influence of climate change on insect development will be briefly outlined.

2.1 Introduction to pest forecasting

Pest forecasting as a concept is important due to the damage that pests inflict. In an agricultural context, pathogens, nematodes, weeds, vertebrates as well as insects and mites are largely regarded all as pests. They are a major constraint to crop productivity and profitability around the world caused by direct and indirect damage to valuable crops (Olatinwo and Hoogenboom, 2014). In the early '90s, such pests accounted for an estimated 45% of pre- and post- harvest (storage) losses worldwide (Oerke, 2007), in addition to losses caused by vertebrate pests (Strand, 2000). Even today with improved mechanisation and crop technology, between 26 to 40 % of the world's potential crop production is lost annually due to weeds, pests and diseases, and these losses could double without the use of modern crop protection methods (FAO, 2012). The yield losses from different categories of pests are estimated to be US$500 Billion worldwide (Oerke *et al.*, 1994; Liebhold and Tobin, 2008). At a country level, the damage can be severe and managing pests can have multiple impacts. For example, accounting for direct damage to crops by the pest, the cost of chemical intervention and potential ramifications in terms of health (i.e., humans poisoned in areas where there is unrestricted spraying), the total economic losses are approximately US$ 17.7 Billion (Oliveria *et al.*, 2017).

The damage caused by pest organisms of any kind is one of the most important factors in the reduced productivity of any crop plant species (Fitt and Wilson, 2012; Metcalf, 1996).

It is hard to address pest forecasting without including the importance of the origins of IPM. We have come a long way from when in ancient times, humans lived with crop pests without major efforts to control them, but as competition increased for food, humans developed ways to safeguard their crops against pests by using management techniques such as cultural, mechanical, physical, and biological control (Abroi and Shankar, 2012). This formed the foundations of the principles of IPM we know today. The use of natural biological control is almost as old as the history of agriculture itself. Chinese cave paintings clearly show ducks being used to consume pests off crops, a technique still operating in rural China in the late 20th Century (Van Emden, 1989). The mid-20th Century was significant with the advent of synthetic organic insecticides which initially provided successful results against pests (Metcalf, 1980). However, as time progressed evidence has accumulated of environmental damage in terms of pollution and pest resurgence particularly in scientific research and publications such as Rachel Carson's Silent Spring (Carson; 1962; Dunlap, 2008). Since their advent, dependence on chemical insecticides has resulted in issues of developing resistance genes in over 500 insect pests by 1990 (Georghiou, 1990).

IPM as we know it today really came into effect in the 1960s with the aim of reducing harmful consequences of chemical pesticides and the need for widespread chemical intervention in the control of pests and diseases (Fitt and Wilson, 2012). Within any IPM approach, monitoring for pests is a fundamental first step in creating a management programme to collect useful and informative biological data on pest species. Insect pests are monitored through a variety of monitoring tools such as pheromone traps, light traps, emergence traps (as displayed in the pine weevil trial data at the centre of this research), pitfall traps and suction traps (Prasad and Prabhakar, 2012). Monitoring data also serves several purposes. These include ecological impact studies (Pathak, 1968, Hirao *et al*., 2008), tracking insect migration (Drake *et al.*, 2002), timing of pest outbreak into agricultural and forest ecosystems (Teagasc, 2020; Klueken *et al.*, 2009), initiating field sampling procedures (Dillon *et al*., 2006; Prasad and Prabhakar, 2012) and timing of pesticide applications (Teagasc, 2020; Merril *et al.*, 2010). This is just an overview of some of the practical methods adopted to aid pest monitoring and forecasting in a working IPM system.

2.2 Relationship between temperature and development/phenology

Insects are incapable of internal temperature regulation and hence their development depends on the temperature to which they are exposed. Temperature is a fundamental driver of insect development and is one of the principal factors delimitating survival and reproduction. Temperature has been identified as the dominant abiotic factor directly affecting herbivorous insects (Bale *et al.*, 2002). Temperature extremes are a cause of significant natural mortality in populations. Warm and cold extremes can offer potential for the development of environmentally safe pest management strategies (Hallman and Denlinger, 1998). Both heat and cold (i.e., thermal stresses) have been used to supress pests since the beginning of insect control and have been often classified under the broad category of physical controls which Metcalf *et al*. (1962) described as methods which employ abiological properties of the environment to the detriment of pests. While the damaging effects of high temperature are more obvious in terms of susceptibility to infection and lost vitality (Eliott et al., 2002; 2005), insects may, in certain cases, exploit high temperatures for their own benefit. Insects infected with viruses, bacteria or parasitoids frequently seek high temperatures to rid themselves of infection and promote their own survival (Heinrich, 1993; De roode and Lefevre,2012). Insects can also exploit low temperatures for their own benefit. Bumble bees for instance carrying a heavy load of conopid parasitoids stay away from the colony on cool nights and expose themselves to low temperatures, thus retarding the growth of the parasitoids and reducing the chances of successful parasitoid development (Muller and Schmid-Hempel, 1993). Sensitivity of pests to climatic factors such as precipitation and temperature varies by species. The direct effects of temperature on insects include effects on developmental traits, in respect to their existing habitats and life histories and crucially their ability to adapt during their ongoing development (Bale, *et al.*, 2002; Regniere *et al.*, 2012).

Temperature is the main driver of key life cycle functions of insects, not only in terms of reproduction but also for movement, development, and survival. Temperature affects life-cycle duration, rate of development, voltinism, population density, size, genetic composition, extent of host plant exploitation as well as local and geographical distribution inked to colonisation and extinction (Bale *et al.*, 2002). Temperature and infection are two of the most common factors exerting selective and targeted pressure on all organisms including insects (Johnston and Bennett, 1996; Thomas and Blanford, 2003; Wojda, 2017). Temperature shapes the processes and outcomes of characteristics that happen throughout life cycle development of insects including during mating interactions and reproduction (Regniere *et al.*, 2012).

The photoperiod and temperature have been shown to have a direct effect inducing diapause (Velarde *et al*., 2002). Tolerance to high temperatures, known as heat tolerance, is a key determinant of an individual insect's survival in warming environments (Colinet *et al*., 2015). Insects often must use physiological mechanisms of heat tolerance and thermoregulation as displayed in dispersal behavior to find cooler sites, and on the presence of refuges with appropriate microclimates to regulate their body temperature (Sunday *et al.,* 2014; Duffy *et al.*, 2015). In recent times, focus has expanded to the metabolic theory of ecology (MTE) where temperature and body weight are the fundamental determinants of the rates at which life's central processes occur including metabolism, development or individual growth, species diversity and even ecosystem processes (Brown *et al*., 2004).

2.2.1 Life cycle and development (Coleoptera)

Insects are poikilotherms meaning they cannot regulate their body temperature outside of behavioural means such as basking or burrowing (Regniere *et al.*, 2012). Early research on insect responses to temperature dates to the 18th century and the work of famed French entomologist René de Reaumur on insect parasitoids and on caterpillars (Carton, 2005). The production of heat by endothermy and the elevation of body temperature have been known for a long time in insects (Himmer, 1932; Krogh and Zuethen, 1941). More recent evidence suggests that flight activity in adults, as well as rate of growth of the immatures in social insects, are sometimes dependent on elaborate mechanisms of temperature regulation (Heinrich, 1973). Thermal limits of insects have been studied as far back as the early 1900s, including factors such as the relationship between insect development and temperature, identifying factors such as cold resistance and the influence of relative humidity as well as their importance in respect to insect phenology (Bacmetjew, 1900; Payne, 1929). These studies identified important drivers or regulators of insect response to temperature such as the role of extreme temperatures of differing intensity and quantity, the importance of cofactors such as relative humidity and finally the critical importance of individual variation within a population (Rebaudo and Rabhi, 2018). Temperatures that exceed specific optimum ranges for an insect can lead to decreased growth rates, reduced fecundity and increased rates of mortality for many species (Friendenberg *et al*., 2008).

Expanding and optimising the use of laboratory and field work data coupled with parameterising the thermal response of insects in particular and poikilotherms in general will enhance the development of process based phenological models to encompass such characteristics (Regniere *et al*., 2012). This measure of accumulated heat energy is known as physiological time.

Physiological time provides a common reference for the development of organisms during their life cycle and represents an important factor in biological development (Trudgill et al 2005).

There are three important effects of temperature on insects which are closely interrelated to insect performance. Firstly, temperature affects rate assessed processes such as development, reproduction, and movement. Secondly, acute, or chronic exposure to low or high lethal levels inflicts mortality. Thirdly and finally, there is a range of sub-lethal (delayed) effects that can lead to abnormalities to the normal response to temperature within the favourable zone of inhabitation (Van Emden and Harrington, 2007). The effect of temperature on development in insects is often displayed using the inverse of development time, that is, the development rate. The development rate of insects starts from a critical thermal minimum (CT_{min}) and increases slowly as temperature increases. It reaches a temperature range where development rate is almost linear, then continues growing up to an optimal (T_{opt}), to finally decrease rapidly to a critical thermal maximum (CT_{max}) (Rebaudo and Rabhi, 2018). It is also worth noting the variability in terms of developmental rates amongst organisms and how these can vary within populations as well as influence the observed distribution of phenological events in those populations (Yurk and Powell, 2010). For many ectotherms, the relationships between ambient temperature and development, survival and reproduction scale up from daily or even hourly effects on individuals to seasonal patterns of phenology (Visser and Both, 2005). As poikilothermic organisms, the body temperature of insects can vary considerably. They mainly rely on ambient temperature and must adapt to survive any arising environmental stress.

The metabolism of all insects is affected by environmental conditions. The influence of the environment can limit an insect's development and its ability to remain active, and ultimately, its survival (Dent, 1991). The effect of environmental extremes has often been regarded as largely an all or nothing response, so that is why marked activity thresholds exist and can be important in relation to temperature dependent phenological models such as flight, etc. (Taylor, 1963). Low temperature can be a potential lethal stressor for all insects. There has been intensive research into the effects of temperatures on winter survival and the subsequent spring abundance and pest outbreaks including the possible effects of warming climates (Bale, 1999). From a beetle (Coleoptera) perspective, it is worth noting that research has shown winter warming as the primary cause for beetle declines in terms of both abundance and diversity (Harris *et al*., 2019). Laboratory methods and field trials have shown that coleopteran species are negatively impacted when exposed to extreme temperatures. One study exposed the red flour beetle (*Tribolium castaneum),* a serious pest of stored food products, to five constant temperatures from 20 to 35 degrees C showing temperature negatively impacting the development and survival of the beetle (Skourti *et al*., 2019). Moisture in the form of relative humidity has also been shown to affect coleopteran development rates (Zhou *et al*., 2010). The development of insects and when their damaging stages emerge is of significance in entomology and agriculture when considering pest outbreaks. It is an essential component for building phenological models used to predict the effects of temperature on development (i.e., temperature dependent modelling).

2.2.2 Linear and Non-Linear models

Understanding life cycles of forest pest like pine weevil is important in terms of understanding how certain components impact their life cycles and the elements that feed phenological models. Linear and Non-Linear models discussed here are vital in terms of establishing relationships between essential components that produce forecast development models. The most common development rate model, referred to as degree-day model, assumes a linear relationship between developmental rate and temperature, between lower and upper developmental thresholds (Allen, 1976). A significant number of models deal with mean development rate versus temperature and the distribution of development times (Wagner *et al.*, 1984; 1985). Simulation models based on mathematical descriptions of biological data such as development data as influenced by the environment are more easily applied across locations and environments to aid pest forecasting (Prasad and Prabhakar, 2012).

Models or other prediction schemes can be utilised to analyse the most effective management actions, based on acceptable control, sustainability, and assessment of economic or other risks (Strand, 2000). Whether simple or complex, a disease or pest model requires essential environmental variables or covariates as inputs, depending on individual pest sensitivity to different environmental factors. Temperature has particular importance as it drives the life cycle progress and development rate of insects. Access to accurate weather data, including temperature, rainfall, humidity, and other relevant measurements, is essential for developing, testing, and evaluating the model performance (Olatinwo and Hoogenboom, 2014). For example, models that have been used for insect phenology, using derived variables from degree-days accumulation, are applicable in most environments, since they utilize knowledge about individual pest species and its sensitivity to baseline temperature that correlates with pest population growth rates (Dawidziuk *et al.,* 2012). The rate of growth and phenological development of an individual plant and insect species has been found to increase almost linearly from a base to an upper limiting temperature threshold (Cesaraccio *et al.*, 2001). The amount of heat required to complete a given organism’s development does not vary.

Physiological time is measured in degree days (DD). One degree day is equal to one degree above the lower developmental threshold over 24 hours (Zalom *et al.*, 1983). Each species requires a defined number of degree days to complete its development. Therefore, the starting date is very important in terms of phenological modelling (Inward *et al.*, 2012; Prasad and Prabhakar, 2012). The date to begin accumulating degree-days is known as the biofix and varies by species. Biofix points are usually based on specific biological events such as planting dates, first trap catches or first occurrence or emergence of a pest (Zalom *et al*., 1983). Population and development models that incorporate developmental thresholds and development rates based on degree days can help growers and pest control advisors to pinpoint biological events. Biological data can also help modelling of later generations based on size of earlier populations (Zalucki and Furlong, 2005). Thresholds are also important in terms of modelling processes. The lower developmental threshold for a species is the temperature at and below which development stops. The rate of development is simply the proportion of development that occurs at a specific temperature over a period of an hour or a day. The upper development thresholds are less clear, but it is often taken as the temperature at and above which the rate of growth begins to decrease (Zalom *et al.*, 1983).

Non-linear methods have also been developed for use in life cycle development modelling. As increasing temperature above a threshold act to reduce development and may ultimately be lethal to organisms, non-linear models are required to account for aspects in insect seasonality at certain temperature extremes (Prasad and Prabhakar, 2012). Non-linear development rate functions based on enzyme kinetics were developed to describe high temperature (Johnson and Lewin, 1946) and low-temperature (Hultin, 1955) inhibition, as well as for both extremes (Sharp and DeMichele, 1977). Various nonlinear models have been developed to describe the temperature response of developmental processes in plants and insects (Yin *et al.*, 1995). Linear models have been recognised as efficient modelling functions within a restrictive temperature range (Campbell *et al.*, 1974; Honek, 1999; Pollard, 2020). More realistic approaches take account of the non-linear, unimodal nature of physiological responses to temperature using the rate summation paradigm which can account for the intrinsic variation of development rates within populations (Sharp and DeMichelle, 1977; Lactin *et al*., 1995). Another non-linear model of temperature dependent development (Stinner *et al*., 1974) utilized a function that is a simple sigmoid curve with an inverted relationship when the temperature reaches above the optimum. This model, as originally given, assumed symmetry about the optimum temperature but can easily be adjusted for asymmetry. The non-linear model by Logan *et al.* (1976) uses an equation that is asymmetric about the optimum but becomes negative for very high temperatures. The interaction of cyclical temperatures with non-linear development can introduce significant deviations from the linear development rate model, especially in the low- and high-temperature regions of the development rate function of such modelling (Worner, 1992). A practical example of how some of these models fit can be shown in the case of the aforementioned (Stinner *et al*., 1974) model which gave the best fit for the Russian wheat aphid developmental rate data as judged by mean square error and successful convergence between observed and simulated data when 14 insects developmental models were tested (Ma and Bechinski, 2008) using appropriate population model design software (Logan and Weber, 1989).

2.3 Development and application of pest forecast models.

In the last 30 years, much effort has been spent on developing pest management systems in terms of first investigating the contribution that might come from ecological tools of life table studies, system analysis and mathematical modelling. For example, many life cycle simulation models have been developed for pest species such as aphids and pine weevil (Van Emden, 1989; Wainhouse *et al*., 2014; Duffy *et al*., 2017). In theory, it was thought that if the role of various factors which cause changes in insect abundance can be understood and properly identified, then a model of this system would enable predictions or forecasts of the consequences i.e., damage of a pest. In pest forecasting, several intrinsic attributes of the pest insect and the determining environmental and host factors need to be considered. Most pest forecast models consider the phenology of the herbivore and its host. Near real time or observed pest incidence data compiled with remote sensing or geographic information system (GIS) tools can often facilitate early warning of impending pest infestation in a temporal and spatial perspective. In addition, the collection and analysis of reliable weather data from pest affected areas is an essential input for predictive models (Prasad and Prabhakar, 2012). Forecasting systems indeed have become increasingly sophisticated also. Even the simplest approaches are based on statistical models, whereas the more sophisticated have required computer models and advanced specialist software (Dent, 2000; Olatinwo and Hoogenboom, 2014). Pest modelling seeks to predict the activity of biotic agents which could adversely affect a crop or plant species, to facilitate the implementation of preventive measures to reduce the potential negative impacts of the pest on a plant or crop.

2.3.1 Deployment of pest models, types and current IPM

Detecting an impending disease outbreak or pest attack early enough to prevent severe economic losses serves as a strong management incentive to utilize support tools where available (Olatinwo and Hoogenboom, 2014). Applied modelling of crop diseases and pests has mostly targeted the development of support capabilities to schedule scouting or pesticide applications (Donatelli *et al.*, 2017). Crop and pest models have many current and potential uses for answering certain questions in research, crop management and policy. Models can assist in the synthesis of research understanding about the interactions of insect physiology and their environments (Boote *et al.*, 1996). Several reviews have documented recent advances made in the field of designing generic simulation models for pests, diseases, and crop losses (Savary *et al.*, 2006; Esker *et al.*, 2012). Process based modelling has been shown to be critical to quantitatively address questions pertaining to the behaviour of complex systems, such as crop pest and pathogen systems (Donatelli *et al.*, 2017). Within the context of forecasting, models seek to quantify the important relationships between the environment and the development of the pest to enable prediction of likely incidence of pest emergence at some future time. These models commonly involve regression analysis of insect number of one stage against another (e.g., adults vs. larvae) or against damage (Dent, 2000). Some of the models most utilised in IPM are the classical linear regression models, for which one may construct ANOVA (Analysis of Variance) tables to study the effects of predictors. Other commonly used models include crop-growth and crop loss models (Koul *et al.*, 2004), economic threshold and injury level models, sampling models and phenology models. In terms of phenological modelling use, a variety of modelling approaches and model development tools have been established that relate temperature to insect development rates or times (Sharp and DiMichele, 1977; Schoolfield *et al.*, 1981; Wagner *et al.*, 1984; Sporleader and Chavez, 2009).

2.3.2 Modelling applications in Ireland and UK

In terms of application of pest modelling in Ireland and the UK, many models have focused on process-based simulation models that aim to enhance knowledge-based decision making regarding both the timing and extent to which insecticides are relied upon in an agronomical context (Duffy *et al.*, 2017). Models developed based on a simple technique of degree-days, may utilize air or soil temperatures to describe the phenology of pest species, to determine when they reach a pre-determined population threshold that warrants pest management actions (Olatinwo and Hoogenboom, 2014). The information may also be useful for more efficient scheduling of pesticide applications based on the known biology of the pest. It has been noted that the degree-day technique has been useful for predicting insect pest populations elsewhere too, such as the European corn borer, rice water weevil and pink bollworm, particularly in tree, vegetable, and field crops, where pesticide applications may be accurately timed using phenology models (Strand, 2000). System models or other prediction schemes can be utilised with appropriate biological, environmental, economic, or other inputs to analyse the most effective management actions in general, based on acceptable control, sustainability, and assessment of economic or other risks (Prasad and Prabhakar, 2012).

A lot of phenological models are based on high quality biological data that is often gathered over several years of research as for example with the recent evaluation of the potato-blight forecasting model identifying the "Irish Rules" forecasting model for the disease dating back to the 1970's developing with further knowledge as time progressed (Cucak *et al.*., 2019). The process-based modelling of the dynamics of plant pests and diseases aims at reproducing the biophysical processes guiding their development and spread in time. The effect of weather conditions has traditionally been an important focus of these models. The dependency of pest growth rates on the variability of weather conditions implies that models should reproduce these relationships by modulating their responses accordingly (Donatelli *et al.*, 2017). Other models of pests in Ireland particularly relevant in terms of temperature dependence include work on the grain aphid (*Sitobian avenae*) looking at the size and timing of populations in response to temperature as well as work on the blue willow beetle (*Phratora vulgatissima)* creating a voltinism/phenological model as a serious pest of short rotation crop (SRC) willow in Europe (Duffy *et al.*., 2017; Pollard *et al.*., 2020). A simulation model in the UK followed a generation of pine weevils through the life cycle to predict voltinism based on temperature in the root stump microhabitat modelled from air temperature. The model also investigated the potential effect of climate change on voltinism through future projections simulated from the UK climate projections 2009 weather generator for the 2030s to 2070s (Wainhouse *et al.*, 2014). This model is explained in detail below as it is providing the core model structure that is adapted for the development of the pineR forecasting tool, outlined in this research.

2.3.3 Benefits and limitations of pest modelling

In terms of certain pests, there can be many benefits from modelling. Forecasts can be derived for optimum flight periods for instance from daily synoptic weather forecasts to facilitate the detection of pest infestations and disease vectors as well as the timing of pesticide applications to intercept and eliminate pest outbreaks during displacement from breeding zones (Das *et al.*, 2007). Quantifying the impacts of plant pests and diseases on crop performances represents one of the most important research questions that agricultural simulation modelling aims to address (Newman *et al.*, 2003, Esker *et al.*, 2012, Donatelli *et al.*, 2017). A carefully evaluated pest model coupled with weather forecasting output can provide an approach for routine spatiotemporal predictions of potential threats for many diseases of valuable crops, especially those for which IPM can play an important role in the long term (Olatinwo and Hoogenboom, 2014). Several crop growths and phenological models of pests have been used to evaluate consequences of global climate change; including elevated CO_2, increased temperature, and altered rainfall patterns (Boote *et al.*, 1996). Simulation approaches through modelling life cycle development of a pest for instance offer flexibility for testing, refinement, sensitivity analysis as well as field validation of developed models over a wide range of environmental conditions (Prasad and Prabhakar, 2012). Modelling can be a major benefit to predicting pest outbreaks or infestations, but it is also true that model developers must be forthright when describing their model identifying its direct purpose and capabilities as well as its limitations. To improve *Helicoverpa* management in Australia, a comprehensive population dynamics model was developed, which incorporates the spatial structure of the habitat and pest population and explicitly simulates the adult movement within a regional cropping system (Fitt *et al.*, 1995) showing the importance of spatial dynamics, often overlooked in pest models.

Simple growth models are often easy to comprehend and require fewer inputs, are easier to use and widely apply (Ritchie, 1989). Model use can be frequently limited by the unavailability of accurate input data. A lot of temperature models require comprehensive weather data such as daily maximum and minimum temperature data and errors in input data, inadequate calibration of recording equipment and sampling errors are common occurrences (Boote *et al.*, 1996). Calibration and adjustment of certain model parameters or relationships may be necessary to make the model work for a particular site or sites. Validation determines whether the model works with independent data sets; that is, does it accurately predict growth, yield, and processes outside of the training data. Validation does not necessarily offer a full validation as validation is a continuous process hence a lot of models are referred to historically as partially validated (Oreskes *et al.*, 1994). Models also must be tested in diverse environments if modelers wish to make applications, changes or extrapolations to those environments and situations (Boote *et al.*, 1996). In this way, modelling can be limited depending on its specific requirements. Simulation models on the other hand based on statistical theory and mathematical descriptions of biological data as influenced by a given environment (i.e., temperature dependent models) are more easily applied across different locations and environments. Pest simulation models offer flexibility for testing, refinement through correction methods, sensitivity analysis as well as field validation of developed models over a wide range of environmental conditions (Prasad and Prabhakar, 2012).

2.4 Influence of regional specific factors on pest modelling

The issue of scales and regional specific factors offers potential insights for modelling as insect pests are traditionally natural disturbance agents themselves that can significantly alter the structure and composition of forested landscapes and thus can impede their ability to provide critical ecosystem services (Munro *et al*., 2022). The quantification of the spatial dynamics of pest populations especially in relation to associated environmental changes (e.g., wind direction or rainfall) is not a new practice. Geographic Information Systems (GIS) utilize site coordinates to store information such as altitude, temperature, soil type and distribution of crop varieties on a spatial map layer (Dent, 2000). Other factors affecting pest distribution should be investigated such as presence of natural enemies or evidence of activities potentially contributing to a pest problem such as with high disturbance in agriculture (Landis *et al*., 2000). With any insect pest, the frequency of monitoring is determined by the biology of the pest, but the crop of course is equally relevant. The crop may have a low damage threshold for instance and therefore may need more maintenance. The area of the crop to monitor also varies as it can depend on the crop type, the farm system or forest size and indeed the pest population structure. However, in any monitoring of scale of pest outbreak for example, it is advised that there should be enough of an area monitored to provide an adequate representation of the field. Record keeping is essential for decision making and evaluating trends in pest populations emerging from season to season (Koul *et al*., 2004). It is important to note too how site-specific factors can interact differently with insect pests which themselves act as the biological drivers of habitat (i.e., forest or crop location) disturbances. Often, interactions among disturbance agents or site-specific factors themselves amplifies the disturbance impacts on site beyond that of any single agent or cause, particularly under a warming climate (Raffa *et al*., 2008).

2.4.1 Site specific factors impacting the status of pine weevil as a pest.

The influence of location is significant in terms of pests and specifically pine weevil. It is difficult to predict the size of pine weevil population at any given place (Von Sydow, 1997; Zumr and Stary, 1994). Pine weevil populations on a clear-cut are affected by landscape factors, such as the presence of the older clear-cuts within a certain distance (up to 5km) as a source of weevils (Lyons, Coillte, *Pers. Comm.*, 2021). Based on field studies in Sweden, pine weevil can fly on average 10km with some individuals reaching distances of up to 80km.

Migration has been identified as a major factor in pest populations (Sobreck, 1980). The extent and scale of pine weevil flight is unknown. Data from Sweden and the UK respectively indicate that weevil flight is temperature dependent, only occurring at temperatures greater than 18^{o}C (Solbreck and Gyldberg, 1979) and their ability to fly declines due to regression of flight span muscles after the main migration period in May (Tan *et al*., 2011). In addition, neighbouring forests are also a source of pine weevils, where they exist in low numbers and from which they can travel to infect clearfell sites. Suitability of standing forests as a source of adult weevils varies. A study in the Czech Republic found that old spruce forests had the lowest pine weevil population, compared to pine, and mixed forests (Zumr and Stary, 1994). The species of previous crops can affect weevil populations multiplying on site with weevil developing much faster and in higher numbers in pine than in spruce (Bejer-Petersen, 1975; Dillon and Griffin, 2008). Soil type has also been found to be a significant influence on incurred damage as weevils prefer to move across peat soils as opposed to a mineral substrate (Dillon and Griffin, 2008; Kapranas *et al*., 2017). Indeed, research has shown the amount of damage by weevil infestation was lowest on pure mineral soil and highest on undisturbed humus (Petersson *et al.*, 2005). The impact of vegetation or brash on weevil damage is unclear.; positive, negative, or neutral effects have been reported (Nilsson & Örlander, 1995; Örlander & Nilsson, 1999; Örlander *et al*., 2001; Wallertz *et al*., 2005).

2.5 Influence of temperature on Large Pine Weevil development

Rates of development of eggs, larvae, and pupae of pine weevil increase linearly with temperature such that weevils reared at higher temperature grow faster, but contrary to expectations are also larger in mass (Inward *et al.*, 2012). Over 80% of ectotherms mature at a smaller size when they develop at higher temperatures. This is known as the Temperature-Size rule (TSR) (Angiletta and Dunham, 2003). Pine weevils follow an Inverse TSR increasing with temperature.

Being larger has advantages such as being less susceptible to mortality (i.e., having more fat reserves reducing overwintering mortality), and larger females will be more fecund and lay larger eggs increasing offspring survival (Inward *et al.*, 2012). The influence of temperature in this case is shown to be positive in terms of effects on fecundity and overwintering survival. Warmer temperatures can also lead to an extended feeding period and a faster generation time for weevil development leading to the potential for considerably increased damage (Willoughby *et al*., 2017). The pre-pupal stage is worth noting also as the length of time was highly variable with a non-linear relationship between development time and temperature. However, as temperature decreases, there is a step-like increase in the length of the pre-pupal stage (Inward *et al.*, 2012). An extended pre-pupal stage seems to minimise the risk of overwintering in the pupal stage and synchronises emergence in the spring for adults. The data from Inward *et al.* (2012) was important for the development of the life cycle simulation model for pine weevil described below.

2.6. A life cycle simulation model for pine weevil (Wainhouse et al. 2014)

Wainhouse *et al.* (2014) developed a phenological model to determine geographic voltinism (i.e., number of broods per year) of the large pine weevil (*Hylobius abietis)* in the UK. The simulation model estimates a generation of weevils through its life-cycle development to predict voltinism based on temperature in the root stump microhabitat, modelled from surface air temperature. Daily air temperatures for representative UK locations were simulated using the UK Climate projections 2009 generator (Jones *et al.*, 2009) for the 2030s – 2070s to assess potential effects of climate change on voltinism. Specifically, the model projected the development and timing of seasonal activity of pine weevil as the future climate of the UK warms, estimating how this might affect management of the forest pest, identifying voltinism of large pine weevil in present and future climates according to region. The model code from this research formed the basis of the adapted `pineR` package for forecasting emergence of large pine weevil in Ireland.

2.6.1 Conditions, inputs, and functionality of the UK model (Wainhouse et al, 2014)

Temperature measurements were made at 17 sites across the UK presenting a range of environmental conditions under which pine and spruce trees are planted and exposed to different environmental conditions. Scots and Corsican pine predominated in the south and east and Sitka spruce predominated in the north and west. Experimental sites had been clear-felled up to six months previously and were on level or moderately sloping ground with minimal shading by surrounding mature trees. At each site, air and root stump bark temperatures were measured within an area (6m by 6m) maintained clear of brash and vegetation with minimal disturbance to the soil surface. The temperature of root stump bark was measured on a single representative root-stump at depths of 10 and 30cm below ground level. Site specific details were recorded at each site covering coordinates (longitude and latitude), tree species, time of felling and the cross-section area and height of the root stump projecting above the ground. Local observed rainfall data and where available, sunshine hours were obtained from the nearest meteorological weather station.

Temperatures over the recording period were used to develop a recursive model to predict daily mean root-stump bark temperature from the mean air temperature calculated from the minimum and maximum air temperatures. The future temperature simulations were derived using the weather generator, which creates a synthetic time series of weather variables at 5km resolution. Regional variation in temperature up to the 2070s was assessed via UK wide maps produced by the weather generator.

The UK-wide maps were used as a basis to select two to four representative conifer forest sites in each of the five regions. Using version two of the generator (Jones *et al.*, 2009), daily minimum and maximum air temperatures were simulated for a total of 12 sites, each of which corresponded either to an experimental field site or was within 25 miles of one. Daily air temperatures were estimated for the baseline period (1961-1990), incorporated by default as a "control" in weather generator simulations, and for the 2030s (2020-2049), 2050s (2040-2069) and 2070s (2060-2089) using the medium emissions scenario (A1B) (Jones *et al.*, 2009). The daily mean air temperature was derived as the mean of the maximum and minimum air temperatures in the weather generator data and was used to predict root stump temperature at depths of 10cm and 30cm. The temperature at 20cm was simply calculated as the mean of values at 10cm and 30cm. These microhabitat temperatures were the basis for the modelling of future seasonal development of weevils across the UK sites. Differences in temperature from the baseline period (or model control) to the present were assumed to have had little effect on -development changes relative to that predicted for the climate of the 2030s onwards, and thus this data was referred to as the current climate or control climate in the model.

2.6.2 Predicting seasonal activity and stages in the UK life cycle model.

In terms of predicting weevil seasonal development and activity, the life cycle simulation model was developed to emulate a generation of weevils from egg to adult oviposition. Voltinism was also determined. The model runs through a series of "life cycle" stages identified by 0-10. These stages are grouped into a series of independent modules (as shown in Figure 2.1 below). The stages 0 and 10 refer to the start time of oviposition in the first generation and to the egg selection for the start of subsequent generations. A default population size of 1,000 females was used to characterise variation in weevil development with each individual weevil progressing through all modules. By modelling a fixed population, it was assumed in the original implementation that they were independent of mortality.

For the primary stages in the life cycle, development or activity was described by a set of parameters, based on temperature and/or time in days or day counts. These parameter values were derived largely from laboratory and field trials across northern Europe (Wainhouse *et al.*, 2014); however, some of the parameters were experimentally derived and a number were included in the model to account for observed attributes of the life cycle such as the timing of the transition from prepupal to pupal stage, or to define a specific overwintering period. Threshold temperatures for adult activity and duration of maturation feeding for weevils was less well defined than other earlier stages. For parameters obtained from experimental data, statistical models were fitted to the data and used in the simulation model. The stages of the UK model follow through the modularised approach in order as displayed below in Figure 2.1.

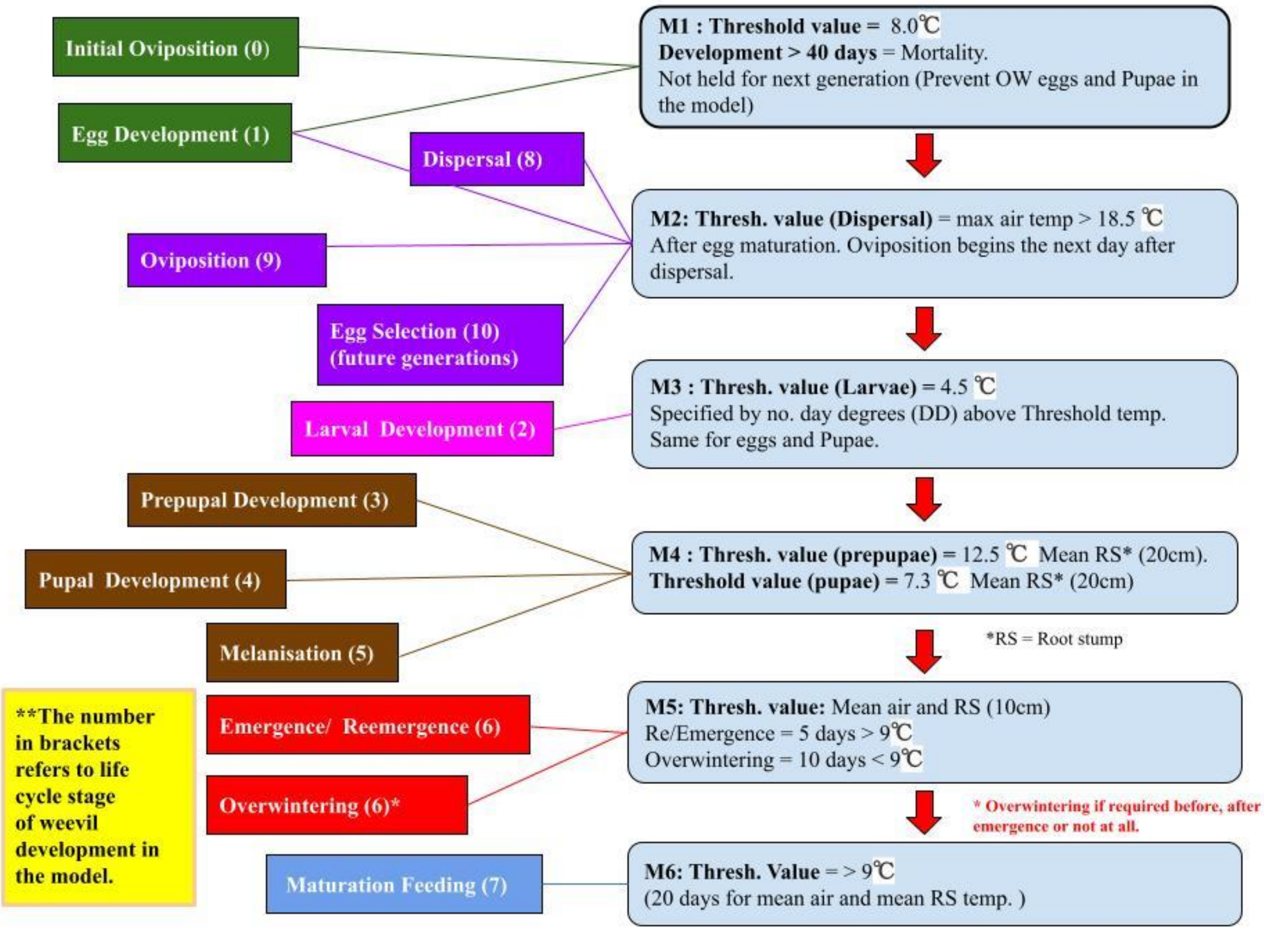


Figure 2.1: The modular approach of the UK life cycle simulation model for large pine weevil *(Hylobius abietis)* (Wainhouse et al., 2014), implemented in pineR.

2.6.3 Life cycle stages of the UK model

The life cycle stages are described below in the following sections according to progression within the model.

- O***viposition and egg development (Stages 0 and 1)***. This is initiated during the Spring oviposition period which is mid-late June (~day 171; 20th June). Moving onto further egg development, the temperature thresholds become very important in the model. The 8 °C threshold for root stump temperature at the 10cm depth is key for development progression. At this stage, maximum development is set at 40 days. Those that take any longer are individuals assumed to have died and are not available for random selection for the start of the second generation run in the model.
- ***Development of Larvae (2), eggs and subsequently pupae*** at different depths are specified by the required number of day degrees (dd) above the developmental threshold temperatures. The threshold for larval development is set at 4.5°C for root stump temperature. Here it is worth noting that 50% of larvae were assumed to develop at 10cm and 50% at 30cm.
- ***Prepupal and pupal development (3 and 4)*** had notable conditions. A facultative diapause results in a nonlinear relationship between development time and temperature, with no obvious developmental threshold temperature. Prepupal development was therefore expressed as the number of days required at a given temperature, based on previous research (Inward *et al*., 2012). It was assumed under varying conditions, the number of days for pupal development was related to mean temperature over this period. Depth selection would only apply to prepupae and subsequent pupal development in the model. The model was run at 20cm as a default depth, but the option could be removed if more convenient to do so. 20cm is the average of the two measured depths of 10cm and 30cm. Regarding the transition from prepupae to pupa, the threshold is identified as 12.5 °C. Pupae taking longer than 60 days are treated as prepupae waiting for the transition threshold to be hit. The threshold value here is 7.3 °C. Geographically, 12.5 °C allows autumn pupal development in Scotland but rapid falls in temperature prevent further development or completion. This would result in overwintering of pupae which does not occur here.

- ***Adult activity and emergence (Stage 6)*** begins with ***melanisation*** (darkening of the beetle) which is indicated at 21 days. Emergence for summer/autumn or pre-winter emergence is indicated with 5 consecutive days (warm days) with mean air and 10cm root stump temperature more than 9 °C. It is presumed this occurs during melanisation or otherwise post melanisation. After emergence of overwintering weevils, they become inactive in further cold spells. This is where the mean air and 10cm root stump temperature is less than 9 °C. Therefore, they do not hibernate with a requirement to re-emerge in the model when it warms up again or passes the threshold temperature.
- ***Overwintering (Stage 6)*** is triggered when the mean and 10cm root stump temperature is less than 9 °C for 10 consecutive days (cold days). ***Re-Emergence (Stage 6)*** is indicated for overwintering previously emerged adults after the first incidence of 5 consecutive warm days with mean air and 10cm root stump temperature more than 9 °C.
- ***Maturation feeding (Stage 7)*** is indicated with 20 days (minimum 12 days) with mean air and 10cm root stump temperature more than 9 °C after overwintering. Where autumn maturation is permitted, setting a notional feeding value (NFV) (varies from 0-1) effectively alters the number of days required as the NFV's changes for different months. If at the start of winter, maturation is incomplete or mature, but dispersal threshold has not been reached, the weevil population overwinters in these criteria. No further accounting is needed before they disperse at the next dispersal threshold or day when reached. Egg maturity days are eligible for accumulation after re-emergence/emergence depending on settings of overwintering requirement in the model. Weevils are inactive on cold days.
- ***Dispersal (Stage 8)*** is indicated when max air temperature hits 18.5 °C after the maturation period is complete.
- ***Oviposition (Stage 9)*** to begin a second run of the model (and subsequent runs) starts the day after dispersal is initiated. It occurs over a total of 78 days when mean air and root stump temperature is more than 9 °C. Oviposition rate declines linearly to zero over this period with a mean of 1.2 eggs/day (egg laying) per individual. If incomplete at the next winter threshold temperature, oviposition is ended and not resumed.

- ***Egg selection* (future generations) (Stage 10)** occurs again with eggs taking longer than 40 days to develop not available for selection and indicating mortality as said. If development of all laid eggs by an individual weevil is less than 40 days, the female reverts to pre dispersal state of maturation female that overwinters where necessary, disperses and oviposits as defined (Wainhouse *et al.*, 2014).

2.6.4. Validation and predictions of UK life cycle simulation model in Ireland (IMPACT project)

Research in Maynooth was undertaken as part of an INTERREG project (IMPACT) (Williams et al., unpublished) used empirical data on weevil emergence sites collected at 4 spruce and 7 pine clear-fell sites (Figure 2.2) and compared these observed data with simulations obtained using the simulation model, developed originally for the UK, and described above. Statistical comparisons showed that for half of the trials the forecast simulations did not differ significantly from observations recorded for weevil emergence. (Williams et al., unpublished). Four sites displayed similar emergence patterns to the results predicted by the life cycle model. One site showed no significant difference with a chi-square test, but a notable difference with the Kolmogorov-Smirnov test. Three sites were significantly different from the life cycle model predictions (Williams *et al.*, unpublished). In addition, four sites were selected at spruce locations across Ireland to investigate voltinism patterns using the UK model with simulated climate data to evaluate potential changes in voltinism due to climate change in Ireland (Williams et al., unpublished)

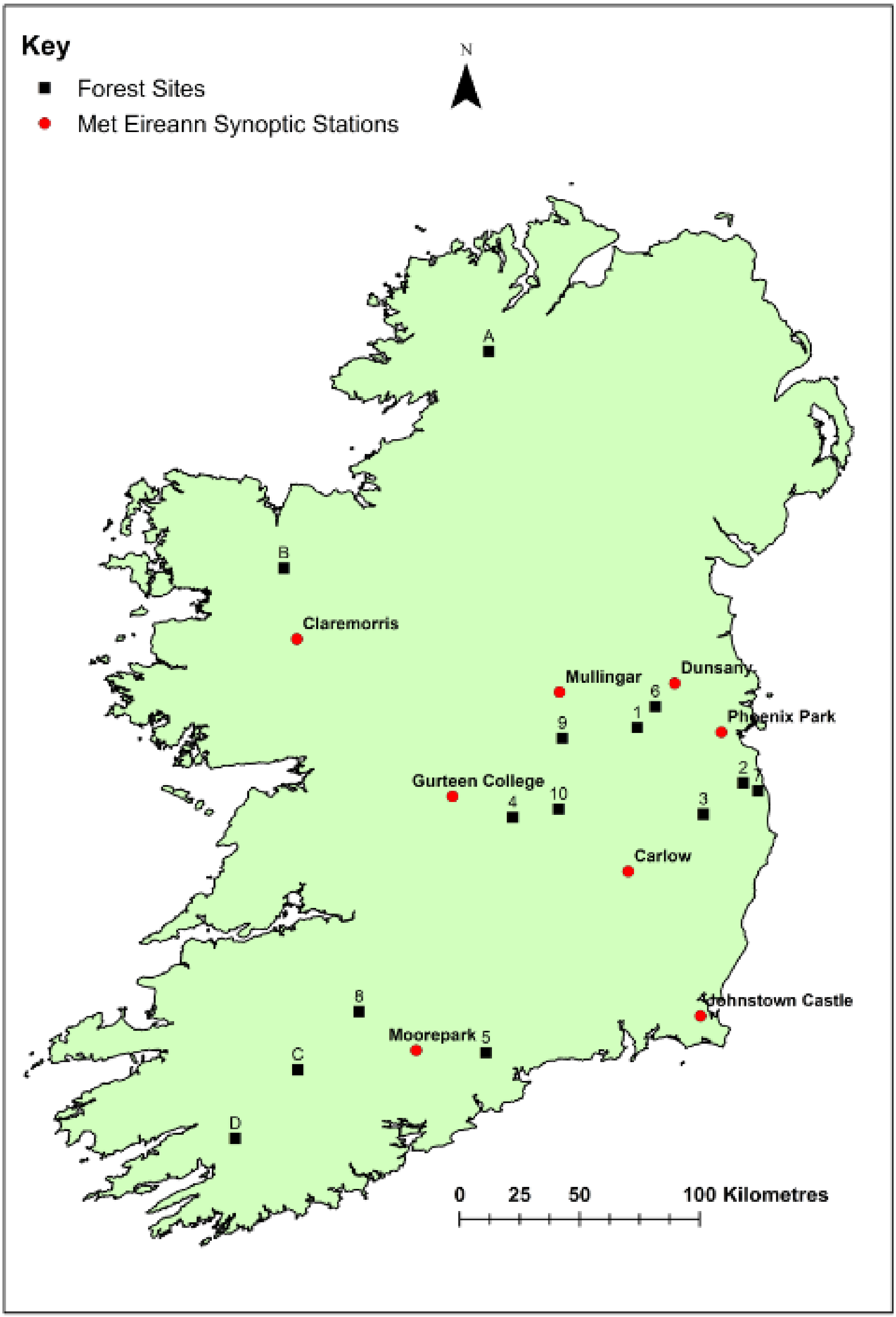


Figure 2.2: Locations of forest trial sites used in the IMPACT project and nearest automatic weather stations (AWS) (Williams et al., unpublished).

2.7 Conclusion

Insect monitoring involves the regular surveillance of key insect traits such as population, biology and movement or migration amongst other criteria. In certain cases, it may be difficult to ascertain what pests are causing damage to a crop. In this case, symptoms of crop damage must be investigated (Dent, 1995). This is important to assess the pest situation and identify pest activity occurring and aid future decisions on a farm or relevant enterprise. The threat of a disease epidemic or pest outbreak is real, hence continuous monitoring is required to avert significant damage to valuable crops from one year to the next. Accurate and reliable weather-based forecasting models can play their role as critical components in IPM strategies. It is not only important to protect the valuable crops and increase productivity but also for the efficient - and diligent use of harmful chemical intervention through pesticides only if essential and overall protection of the environment (Olatinwo and Hoogenboom, 2014). There are certain limitations to pest forecasting models as discussed especially with certain weather parameters but as our knowledge of the biology of an individual pest improves and with the application of emerging technologies, the accuracy of such forecasting models is expected to become a lot more reliable in future enhancing long term IPM strategies and potential to contain harmful pests (Dent, 2000). Pest monitoring is the foundation for the issues for early warning systems, development and validation of pest forecast models and decision support systems, which are crucial for the design and implementation of a successful IPM programme (Prasad and Prabahakar, 2012). The foundations of pest forecasting are equally fundamental to the modelling of pine weevil emergence in Ireland.

Chapter 3: Data and Methods

The modelling framework for this project adapts an established phenological model to predict site-specific patterns of pine weevil emergence based on categorical weevil data and local temperature data. The forecast model builds off an existing UK model (Wainhouse et al, 2014) which was extensively reviewed and tested for this project by the researcher. The forecast model was then adapted to include the effect of site-specific factors to forecasting first emergence of large pine weevil in Ireland. This process of modelling was done via the creation of the `pineR` forecast R package (Lemos Dos Santos et al., 2023).

In this chapter, the methodology to forecast year 1 emergence of pine weevil at 27 trial sites in Ireland using the UK model, as adapted by Lemos dos Santos *et al.* (2023) via the `pineR` forecast R package is described. The data types inputted in the model will first be described, divided into phenological, meteorological and geographical data. The data was compiled by the researcher from existing pine weevil research in Ireland.

Following this, the forecast package (`pineR)` created by (Lemos dos Santos *et al*. 2023`)` and its implementation using temperature data from two sources (meteorological stations and gridded data) will be described.

Site specific factors are then investigated to evaluate the role of site-specific covariates on the forecast bias, identifying factors that may influence pine weevil development rates. A ML (Machine Learning) method is then used to identify these factors and to correct bias in the model and will be briefly outlined.

3.1 Forest Trial Sites

This research was facilitated by a body of existing data on the large pine weevil *(Hylobius abietis)* from nearly 20 years of Irish forestry research. This data encompassed information on insect emergence, population structure and site-specific factors collected during field trials of biocontrol agents undertaken across forest trial sites in Ireland. IPM strategies for weevil management using biocontrol agents such as entomopathogenic nematodes (EPNs) and entomopathogenic fungi (EPF) were investigated in such trials. The trials were conducted by MU researchers and Coillte forestry operatives from 2001 to 2017. Data on the trials were available in the MU Biology Department and in published records (Dillon *et al.*, 2006; 2008, Williams *et al.*, 2013; Kapranas *et al.*, 2017).

In the current research, 35 sites were initially examined before selecting 27 – the reduction in sites was mainly due to lack of appropriate data. The location of the site coordinates was initially verified by access to site maps maintained in the MU Biology Department and are shown in Figure 3.1 and Table 3.1.

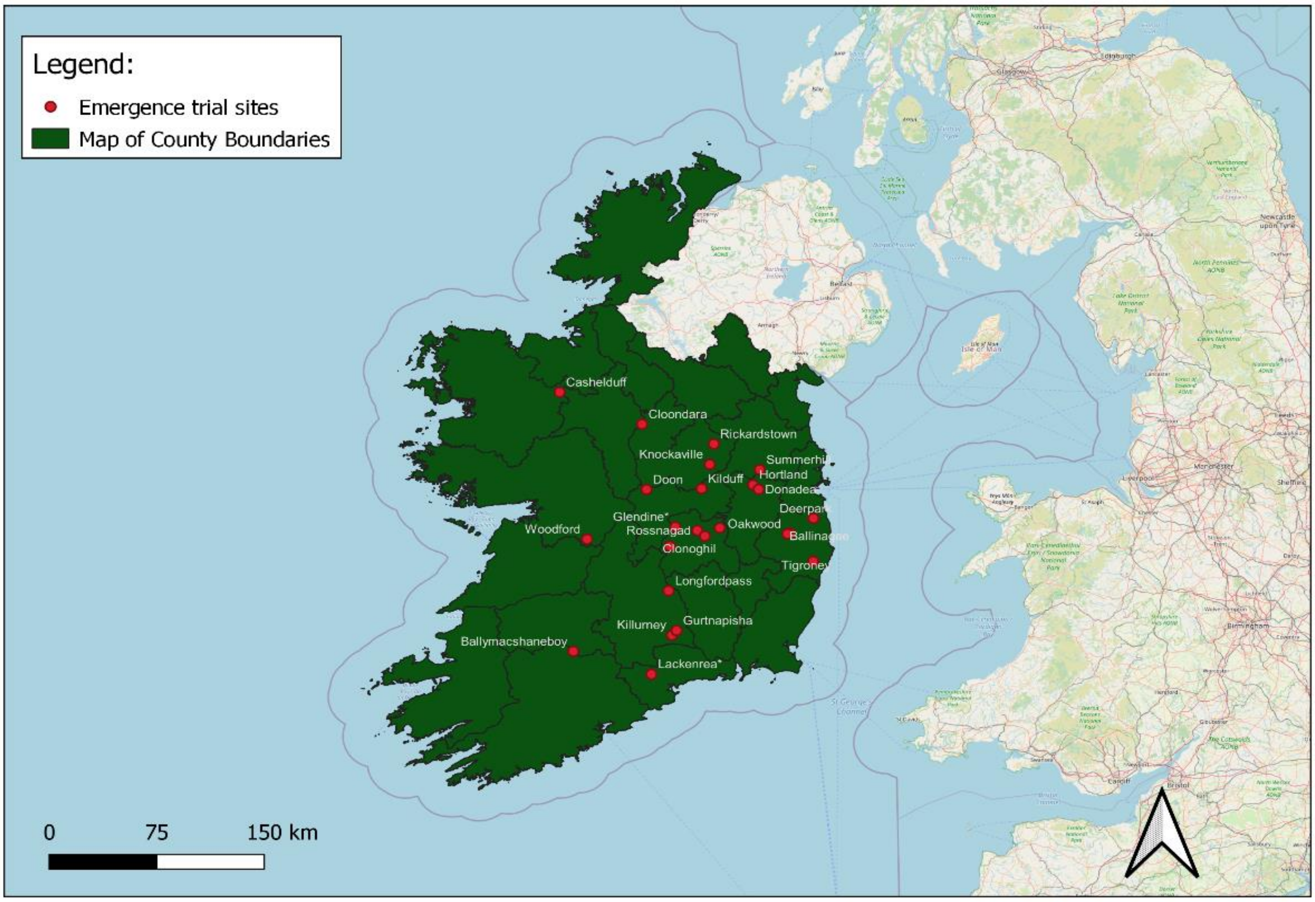


Figure 3.1*:* A map of research trial sites where data was utilized in the modelling of this project to forecast emergence of the large pine weevil (Hylobius abietis) in Ireland. *Sites at Glendine, Lackenrea and Ballyroan have more than one trial plot on site.

Table 3.1: Site specific covariates of Large Pine weevil trial sites examined for incorporation in the modelling to forecast first emergence of the forest pest in Ireland for specific years of trial including slope (angle & %), aspect, elevation, soil type and tree species (Additional site data in appendices Table A2).

Site Name	Tree species	Soil Type	Elevation m	Aspect	Slope Angle degrees	Slope %
Annalecka	Spruce	Mineral	352	N/A	N/A	N/A
Ballinagee	Spruce	Mineral	425	South	12.5	22.17
Ballybrittas	Pine	Peat	68	East	7.13	12.5
Ballymacshaneboy	Pine	Mineral	323	North	6.15	10.77
Ballyroan 1	Pine	Peat	100	South	2.05	3.57
Ballyroan 2	Pine	Peat	100	South	2.05	3.57
Cashelduff	Pine	Peat	215	South	3.41	5.95
Clonoghil	Pine	Peat	124	South	1.69	2.94
Cloondara	Pine	Peat	41	West	2.1	3.66
Corracloon*	Pine	Peat	155	East	6.97	12.23
Corrakyle*	Pine	Peat	245	North	9.47	16.68
Deerpark	Spruce	Mineral	260	North	12.96	23.01
Donadea	Pine	Peat	64	North	2.051	3.58
Doon	Pine	Peat	71	West	3.71	6.48
Emo*	Pine	Peat	92	N/A	N/A	N/A
Glendine	Pine	Peat	380	South	9.35	16.46
Glendine trial 1	Pine	Peat	380	South	9.34	16.46
Glendine trial 2	Pine	Peat	380	South	9.34	16.46
Gurtnapisha	Pine	Mineral	466	North	7.60	13.34
Hortland	Pine	Mineral	76	East	0.41	0.72
Kilduff	Pine	Peat	93	South	3.86	6.75
Killurney	Pine	Mineral	371	South	14.52	25.89
Killnaconnigan*	Pine	Peat	72	N/A	N/A	N/A
Knockaville	Pine	Peat	95	West	4.067	7.11
Knockeen*	Spruce	Peat	65	N/A	N/A	N/A
Lackenrea 1	Pine	Mineral	170	North	6.53	11.44
Lackenrea 2	Spruce	Mineral	175	North	6.53	11.44
Longfordpass	Pine	Mineral	130	West	1.48	2.58
Oakwood	Spruce	Mineral.	250	West	4.77	8.34
Rickardstown	Pine	Mineral	90	East	3.13	5.47
Rossnagad	Pine	Peat	94	West	2.75	4.79
Summerhill	Pine	Peat	90	West	1.17	2.03
Tigroney	Pine	Mineral	246	South	8.94	15.72
Woodford	Pine	Peat	192	South	2.60	4.54

These sites were merely investigated in initial analysis prior but were not included in the forecast model for lack of necessary observed emergence data for appropriate comparison.

3.2 Data

The data outlined covers the key inputs of the forecasting model as well as biological data used for comparison with model outputs, under the headings phenological, geographical and temperature data.

3.2.1 Phenological data

This data refers to the observational data for large pine weevil recorded on trial sites. The biological data includes weevil emergence data that is used to compare against the model outputs, as well as life cycle phases recorded in population structure at stump hacking, data that were not included in the forecast model. Some sites had no stump hacking data, but the final 27 trial sites had the required emergence data.

Site specific data such as altitude and soil type were collected from each of the sites to identify any factors that may aid or provide more reliable and accurate pine weevil predictions.

3.2.1.1 Pine weevil emergence data

Pine weevil emergence data was sourced from 27 sites selected in Ireland. This encompassed empirical data on timing of weevil emergence that was assessed by emergence traps (Figure 3.2). Weevil emergence traps based on a design by Moore (2001) were erected at typically 10-20 stumps per site and emptied at intervals of 2 to 4 weeks, averaging around three weeks. These traps captured emerging pine weevil adults. This data represents the primary phenological data used for simulating or forecasting weevil emergence with the accompanying meteorological inputs in the modelling of `pineR`. Only data from the first year of emergence was used in the present study although year two data was collated for ten sites.

3.2.1.2 Stump hacking data (population structure)

One of the proposed subtasks of the model was incorporating population structure in stumps for emergence trial sites where there was access to prior stump hacking data (17 sites). This data indicates the proportion of weevils at each developmental stage (larva, pupa, and adult) ascertained by the stump hacking on the same sites just prior to the first emergence of adult weevils in the year of the trial. This data is used mostly for comparison purposes at emergence trial sites. However, there is potential for incorporating this data into future research to potentially run trial sites from a specific module in the code or life cycle stage in an adapted iteration of the forecast model. (See discussion chapter).

Figure 3.2: An emergence trap for a large pine weevil research trial in Ireland (Attribution: Apostolos Kapranas, 2017)

Figure 3.3: A Coillte operative Stump hacking a pine stump to check for pine weevil larvae at a clear-fell site.

3.2.2 Geographical data

This data details the specific location data or covariates at each large pine weevil trial site and factors that could be considered in the model for the model to forecast emergence patterns more reliably.

3.2.2.1 Elevation, aspect, and slope

The altitude for most pine weevil trial sites was recorded by field trial researchers. Where it was not recorded, values were generated for site coordinates via a 30m DEM (Digital Elevation Model) for Ireland (ALOS, 2021) and raster analysis in QGIS (V.3.22.0, 2022). Calculations via the raster analysis for site coordinates also calculated slope angle which was converted to percentage via an internal algorithm for clipped raster files over areas of interest. Aspect was also generated via raster analysis with appropriate GIS tools in QGIS. Slope and aspect values were considered relevant as they can potentially affect the amount of solar radiation reaching the soil (McCutchan and Fox, 1986) as well as potentially affecting the microhabitat distribution (relative to soil surface) of developing weevils depending on where the stumps are located. Altitude has been shown to be relevant in research trials with the lowest beetle captures recorded at lower elevations where climate conditions were warmest (Harris *et al.*, 2019). The issue of colder temperatures at higher elevated sites is also of note as cold sensitivity is relevant to insect development as ectotherms and how they respond in such environments (Hallman and Denlinger, 1998; Denlinger and Lee Jnr., 1998). An elevation map of all sites in the model can be seen below in Figure 3.4.

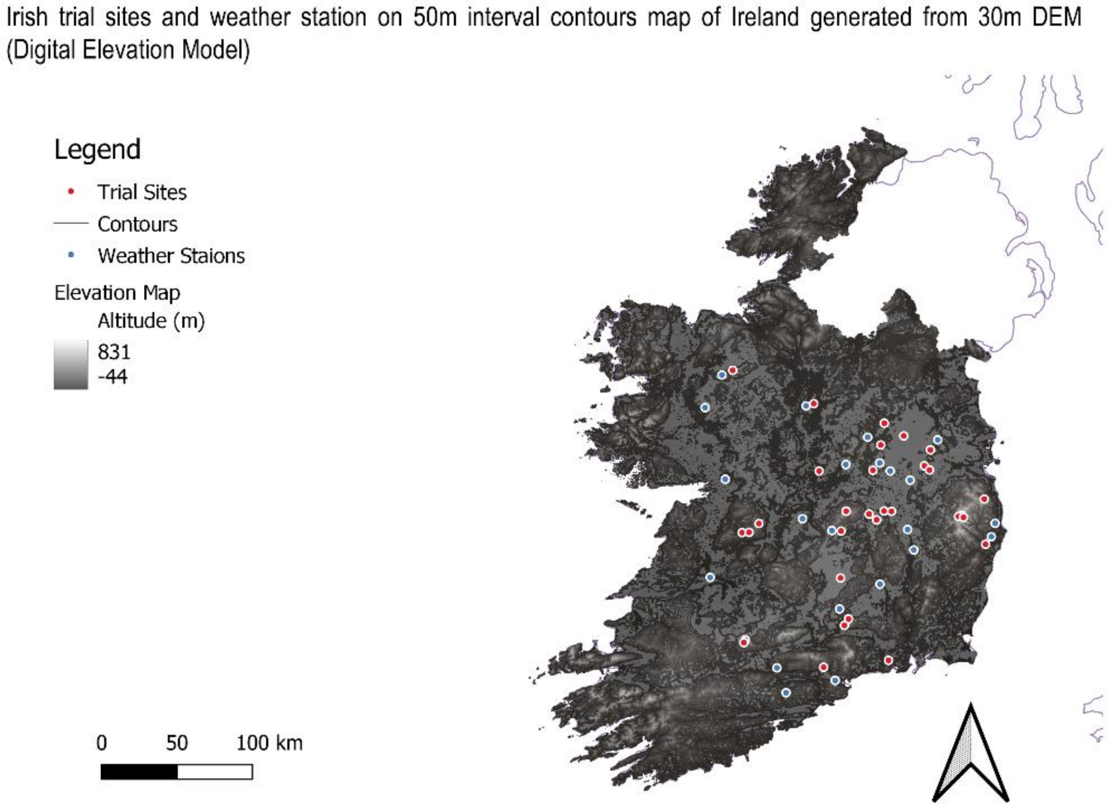


Figure 3.4: An elevation Map of Ireland with contours utilizing a 30m DEM (Digital Elevation Model) displaying pine weevil trial sites and weather station incorporated in the `pineR` model.

3.2.2.2 Soil type

The soil type at each of the trial sites was classified as either peat or mineral soil. This was indicated according to the research trial records and in the few sites where it was not recorded by verification of the location of sites against national soil survey maps and GIS maps (National Soil Survey, 1980; Teagasc GIS, 2021).

3.2.2.3 Tree species

Tree species information was collected from each of the relevant research trials classified as pine or spruce in the model input with the majority of these being pine sites. Tree species refers to the species of trees that were felled, leaving stumps in which pine weevils developed. Of the 27 sites used for the modelling, 24 were pine and 3 were spruce. Although Sitka spruce is the dominant tree species planted in Ireland, pine typically has higher populations of developing pine weevil.

Sites selected for the weevil field trials were largely chosen for the likelihood of infestation of the pest to properly assess the success or failures of biological controls investigated (O'Tuama, *Pers. Comm.*, 2021).

3.2.3 Temperature data

Meteorological data was obtained via historical data from the national meteorological service, Met Éireann. Since there was no suitable temperature data available for the specific sites, the model was validated through two alternative sources of temperature data with two different approaches. The first approach employed temperature data from weather stations located in proximity to pine weevil trial sites. In this approach, a distance weighted temperature dataset was created from the nearest weather stations to weevil trials within a fixed radius for the model. A second approach introduced a gridded climate data product (0.01 degree/ 1km) obtained from Met Eireann (Walsh, 2012).

3.2.3.1 Observed Temperature Data (weather station)

Publicly available observed meteorological data were used to calculate mean daily temperature values (Met.ie, 2021). Both currently functional and closed stations (but functional for weevil trial period) were utilised. The selection of stations was based on data availability and proximity to each weevil trial site. The temperature datasets were compiled to represent the period of weevil emergence at each respective site for the expected life cycle duration and period of observation for biocontrol trials. In total, data from 23 weather stations was used in the model (Figure 3.5, Table 3.2). These stations represented a mixture of automatic weather stations (AWS) (8), manned synoptic weather stations (2) and climatological stations (13) (see Figs 3.6, 3.7).

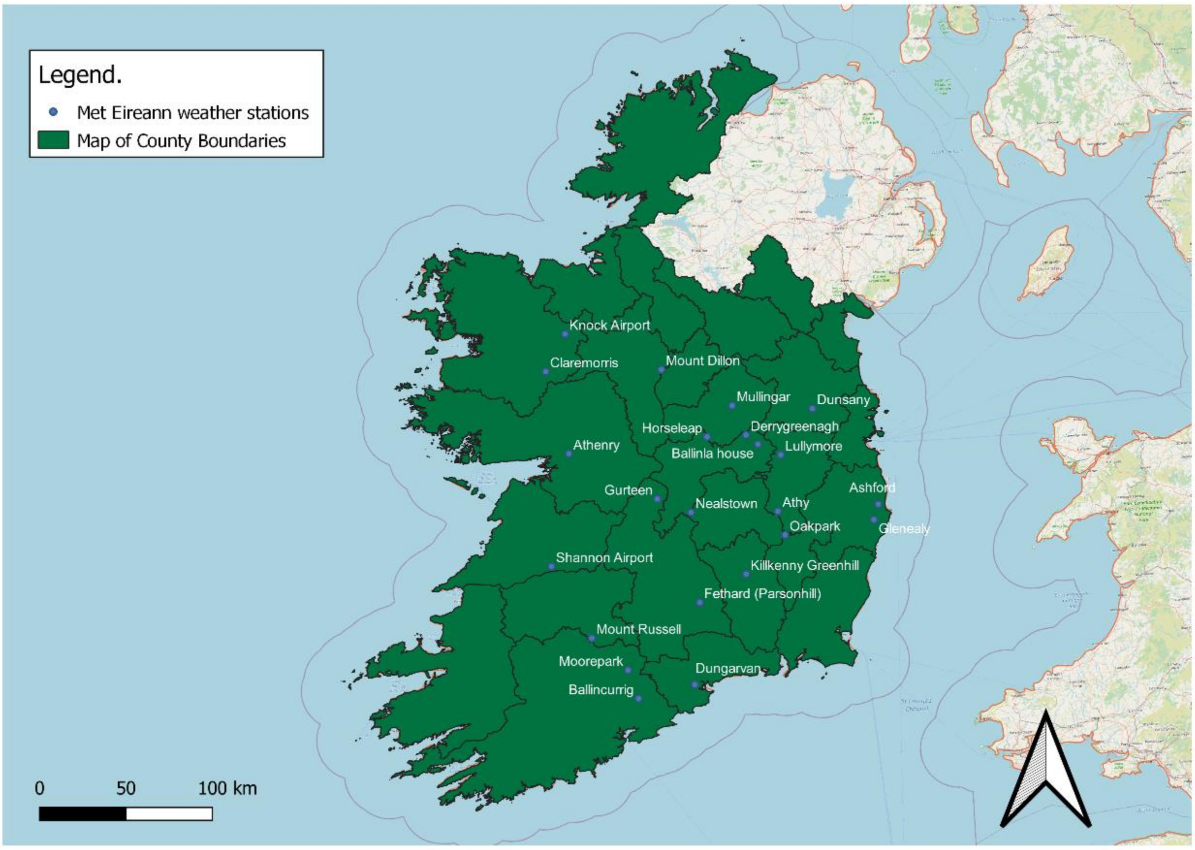


Figure 3.5: A map of weather stations utilized to forecast emergence of the large pine weevil *(Hylobius abietis)* in Ireland.

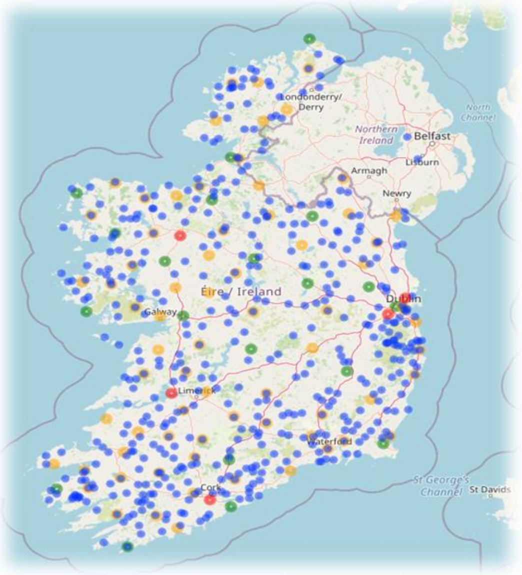


Figure 3.6 (Left) and 3.7 (Right). An automatic weather station in operation, assembled at Teagasc Grange in Dunsany County Meath, where some of the temperature data used in the model was sourced from in accordance with proximity to weevil trial sites in Co. Kildare (Hortland) and Co. Meath (Summerhill) (left) and the national observation network of weather stations in Ireland (WOW.ie, 2022).

Table 3.2: Coordinates of all nearest weather stations incorporated in the modelling in accordance with weighted proximity to weevil trial sites for specific year of trial recorded temperature data and elevations of weather stations.

Station_ID	**Easting**	**Northing**	**Altitude (m)**
Mount Dillon	201448	275143	39
Athenry	147868	226499	40
Oakpark	273036	179492	62
Dunsany	288886	252601	83
Claremorris	134497	273958	68
Moorepark	182130	101304	46
Gurteen	199155	200226	75
Mullingar	242429	254343	101
Dungarvan	220793	92939	18
Derrygreenagh	250412	237271	90
Shannon Airport	137813	161308	15
Knock Airport	145683	295687	201
Ashford	327013	197214	15
Athy	268877	193127	61
Lullymore	270594	225881	85
Glenealy	324494	188337	122
Ballinla house	257329	231932	91
Nealstown	218639	192454	219
Fethard (Parsonhill)	223733	140379	165
Horseleap	227916	236217	72
Mount Russell	161183	119830	195
Ballincurrig (Peafield)	188204	84806	158
Killkenny Greenhill	250546	156852	61

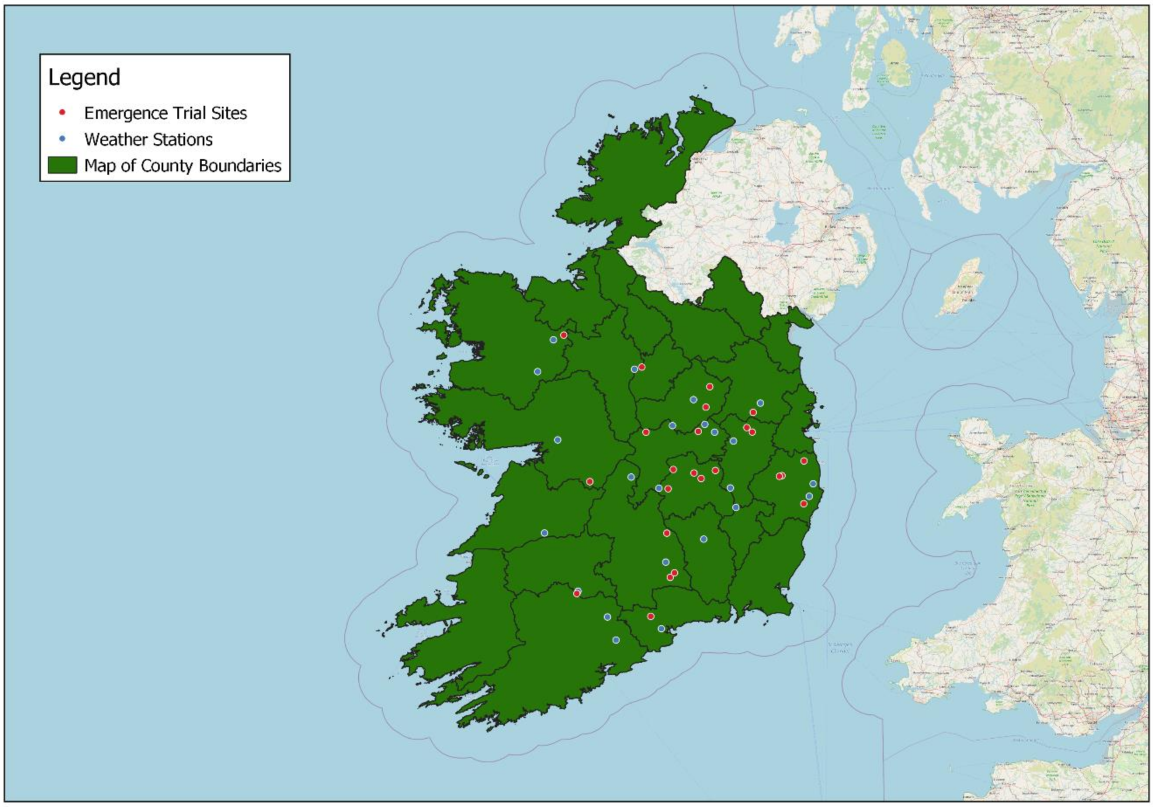


Figure 3.8: All sites utilised in the modelling of `pineR` in Ireland incorporating research trial sites and nearest weather stations.

3.2.3.2 Gridded Climate Dataset

Met Eireann (0.01-degree resolution data) 1km gridded climate dataset (Walsh, 2012) was developed from the observed monitoring network based on spatial interpolation with covariates including elevation. The gridded daily values, at 1 km resolution, explicitly accounts for topography (Flanagan *et al.*, 2019; Walsh *et al.*, 2012). The coordinates for weevil trial sites were used to extract the temperatures for the relevant grids from the Met Eireann gridded data. This was done within R through NetCDF handling libraries to deal with the gridded data to extract the values corresponding to the weighted temperature data derived from surrounding met stations (distance weighted as appropriate) for each site (See Section 3.2.3.1) to compare the outputs of both methods.

3.3 `pineR`: A refined forecasting tool for Ireland

The `pineR` Package developed for this research is an adjusted implementation of the original UK simulation model for voltinism of large pine weevil (Wainhouse *et al.* 2014) (see Section 2.6) (Figure 3.9). For the purposes of this project, extra functionality is implemented with an adapted correction model based on statistical machine learning for refining the estimates of weevil emergence.

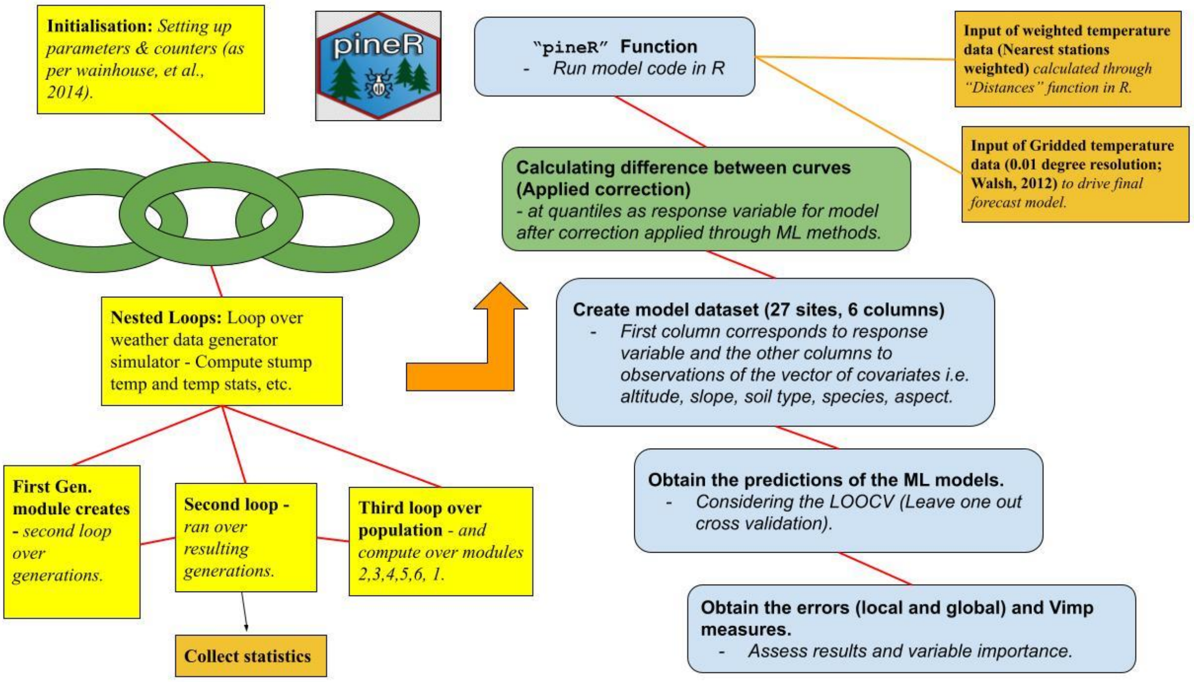


Figure 3.9: The adapted forecast model (`pineR`) enacts a correction through machine learning on the modular approach of the UK life cycle simulation model for large pine weevil *(Hylobius abietis)* (Wainhouse et al., 2014) seen previously in Figure 2.1.

3.3.1 The `pineR` package

The model is structured in three major nested loops which follow an initialization process. After the initial setting up of parameters and counters in the original model, the first loop is initiated. This is where the Wainhouse *et al.* (2014) model runs over the respective input weather data and the model calculates mean daily air temperature, used as input the model. Once the first generation of model weevils completes development, the third nested loop comes into effect. The second loop in between just to clarify is the running over resulting generations of the model. The third loop allows for the simulation of further generations of weevil populations. Loop three works over populations regarding key life cycle stages assigned to 6 modules in a specific order in `pineR`.

These corresponds to dispersal, oviposition, and egg development (2), Larval development (3), Prepupal and pupal development (4), Emergence and overwintering if required (5), Maturation feeding (6) and Initial oviposition and egg development (1) in that order covering the running of the life cycle simulation model. In the construction of the model utilized in this project, it was decided to simplify the original UK model structure with multiple conditions within the code to create a module by module or stage by stage approach, assigning relevant functions and parameters to forecast first year emergence of large pine weevil. An example of the newly assigned functions is the weighted temperature function which was developed for the weighting of temperature data from nearby meteorological stations based on their respective distances from a trail site (distance function). The package and all relevant code and help files are available at https://github.com/Alessandra23/pineR.

The development of eggs, larvae and pupae at different depths was specified by the required number of day degrees (DD) above the developmental threshold temperature for the stages. Individual DD requirements for any stage were generated from normal distributions with the means and standard deviations derived from the original UK model (Inward *et al.*, 2012). For the prepupal stage, development was expressed as the number of day degrees required at a current stage to progress based on previous UK research. The life cycle stages operate as stipulated above in the UK model with the key degree day requirement for adult emergence set at 9^{o}C as a threshold temperature for not only adult emergence, but spring re-emergence, feeding and oviposition as well for the second generation. The first iteration of `pineR` utilized the weighted temperature data from the nearest meteorological stations. The functioning and specific code methodology of the adapted `pineR` model is further explained in the appendices of this thesis.

3.3.2 Weighted averaging of nearest weather station data

The observed temperature data for `pineR` was initially sourced from automatic, manned, or climatological stations that were chosen based mainly on their relative proximity to weevil trial sites (Fig. 3.6). The weighted averaging of this temperature data had specific rules for implementation in `pineR` that led to the creation of the weighted`_temp` function. This worked largely with a `distances` function applied to location of the weather station to organize the weighting of each site and its contribution to the derived 'site' temperature value. The formula for the weighted functioning to generate the temperature datasets is displayed below.

$$\hat{T}_i = \sum_{j=1}^{r} T_{ji} \times \frac{\sum_{k \neq j} d_j}{(n-1)\sum_{j=1}^{r} d_j} \qquad \text{Equation 1}$$

where i refers to the index of the day of the year, j refers to the index of the number of weather stations, T refers the temperature in the weather station (j) and on the day of measurement (i) in the observed temperature datasets, and $\hat{T}_i$ refers to the weighted temperature generated on the day i. Equation (1) above assigns greater influence or weight to sites closer to the trial site location. There is a rule system as to how sites are selected for weighting or not in the model package. Trial sites within a 10km radius of a weather station were considered co-located sites within the model and were used directly as input to the model, without any weighting function applied. For example, the Summerhill site is located 8.2 km from the Dunsany automatic weather station at Teagasc Grange, consequently, temperature data from Dunsany was used directly in the model. There were 8 co-located sites in the model.

Trial sites with weather stations beyond the 10km radius were identified by proximity for utilization in the model via the created `weighted_temp` function in R. Sites beyond 40km were not considered in any weighting function as another rule of the model to reflect as close as possible the temperatures experienced at a given trial site location for the specific year of trial. Weather stations were selected strictly by radius within proximity to trial sites. Stations with significant or mostly missing data for trial periods were not included. Imputation processes were considered in the assembling of data, but missing values were ultimately filled by use of a weighting function effectively treating the missing data as a site itself with the original principles of the `weighted temp` function in `pineR.` An example of this is Lullymore Nature Reserve, which was included due to its proximity to the Hortland trial site, but needed missing data filled by the closest weather data to it via weighting for the relevant year of trial. The blank observations in this case or missing values were imputed from the nearest station data in accordance with proximity to the weevil trial site. The weighting process effectively ordered sites based on proximity for the specific year of observation. The `distances` argument within the weighted temp function acted as a data object to organize weather stations to be weighted. An example for the Hortland trial site is shown in the appendices of this thesis. In most cases, two to three weather stations were selected within a maximum 20km distance difference of the trial site where possible.

3.3.3 Model Validation

The research project devised a core meteorological plan to structure the implementation of temperature in the forecast model. This focused on a two-tier initialization approach to incorporate the two different temperature data sets, including the observed meteorological station data (either co-located or weighted) and the 1km gridded data. The first tier examines the observed weevil emergence data from trial sites matched to model-predicted emergence using observed temperature data from the nearest meteorological station(s). In tier two, the 1km gridded data was employed as input to the model. Differences between model outputs run with the two different temperature data sets used allowed for an assessment of potential systematic biases between the different data sources.

3.3.4 Incorporation of site-specific covariates

The incorporation of site-specific parameters was a critical component in the modelling of `pineR`. In terms of site-specific factors fully factored in the variable importance and cross validation of the model (see table 3.1 above and appendices), certain site data was included in the proposed correction of the model. The site-specific factors included were elevation, aspect, slope (angle & percentage), soil type and tree species (Section 3.2.2). A representation of the elevation values for sites utilized in the modelling is shown in Figure 5.1, where a wide disparity is shown between high and low elevation trial sites but also compared to the nearest weather station utilized in the weighted iteration of the model using observed met station data. It is important to note the elevation of the weather stations employed in the model were not necessarily at the same or similar elevation of the trial sites. In fact, many of the trial sites are in regions of higher elevation compared to the meteorological network of stations. An overall evaluation of dominant factors was taken in the later cross validation to observe effects on weevil life cycle simulation looking at timing of development to see could any influences or mismatches be identified in emergence patterns. This was achieved via appropriate statistical methods through machine learning methods, including extensive work undertaken by a co-modeler on the project (Lemos dos Santos *et al*. 2023). Tree species, soil type, altitude, and slope % were all investigated via linear regression models at all sites to identify variable importance. Eventually, an alternative statistical method was adopted for this process (See below and appendices). Descriptive graphics were created to compare all site-specific covariates of interest for analysis in the results and discussion chapters of the thesis.

3.3.5 Bias correction with machine learning

This section examines how the correction model proposed for first emergence forecasting of large pine weevil was implemented via a machine learning method to correct bias. Cross validation was used to evaluate the machine learning model developed. The observed and simulated emergence data were employed to develop the correction for both the weighted and gridded climate data and is discussed in detail in the Appendices. This includes the consideration of site-specific elements and the adaptation of the existing UK simulation model (Wainhouse *et al.*, 2014) to improve estimates of first year emergence of large pine weevil (Lemos dos Santos *et al.* 2023). Forecasting and classification tasks have been improved by machine learning method algorithms that are applied in many domains of science based on the idea of learning or training such tasks (Refaeilzadeh *et al.*, 2009).

3.3.5.1 Initialisation of correction model

To apply bias correction with machine learning, the initial uncorrected model had to be investigated. When applied to the temperature data, the model outputs displayed a difference between the simulated and observed CDF (Cumulative probability distribution) for the prediction of the year one emergence of large pine weevil. Therefore, to correct this problem in the forecast model, assuming it was a systematic error, a corrective measure was proposed (Lemos dos Santos *et al*. 2023). This correction is outlined as follows.

$$\mathrm{diff}_{jq} = o_{jq} - p_{jq}, \quad \text{Equation 2}$$

where diff_{jq} is as the difference between the observed o_{jq}, and the predicted p_{jq}, day of emergence in site j (j = 1, …, 27), for quantile q.

The work outlined here was undertaken as part of separate, but related research and forms part of the `pineR` package developed by Lemos dos Santos *et al*. (2023) and additional details are provided in the appendices. Therefore, the method is only briefly outlined here. To begin, a Random Forests (RF) model was fitted to the data concerning pine weevil emergence with diff_{jq} as the response. The data used to fit the RF model was collected from the 27 trial sites and contains five covariates of interest. These are altitude (m), slope (angle %), soil type (peat/mineral), aspect (N, S, W, E) and species (pine/spruce). As this is a relatively small dataset in terms of first emergence trial sites, a leave-one-out cross validation (LOOCV) method was used to avoid overfitting. To understand which variables were most important in the forecast model, a variable importance method was used to show the degree to which a single variable impacts on the response value showing the potential effect of the variable on other values (i.e., how slope influences altitude, etc.).

4. Results

In this chapter, the results of the research findings are outlined. The adapted simulation model package `pineR` created to forecast large pine weevil (*Hylobius abietis*) first emergence was run using meteorological data from two different sources. These were 1) the historical weather station temperature data; and 2) the gridded temperature data for areas of interest i.e., weevil trial site locations. Weather station data (co-located or weighted) was used as initial input to the forecast model. Subsequently, the high resolution 0.01 degree (1km) gridded dataset from Met Eireann (Walsh, 2012) was chosen as the input to the final implemented model. Model-simulated emergence data, derived using both temperature datasets, were compared against the observed cumulative emergence recorded at weevil trial sites. The machine-learning corrected version of the model was then run using each of the temperature datasets. Geographical data sourced for each site is analysed as site covariates to evaluate the role of site-specific factors in the simulated results to identify potential variables of importance. The following sections cover the initial exploratory analysis of the data types considered for modelling, the data inputted in the model itself, as well as the final forecast model results.

4.1 Initial examination of model inputs

This section contains an analysis of the phenological and meteorological data prior to implementation in the forecasting model. The observed weevil emergence data provides the core dataset at each trial site against which the model simulated emergence data using the various driving temperature data, namely the meteorological station and gridded data sources, are compared.

4.1.1 Evaluation of phenological data.

Below is an evaluation of the phenological data relevant to the forecast modelling undertook in this research project covering emergence data and population structure data.

4.1.1.1 Emergence data

The cumulative emergence of adult weevils at each of the 27 trial sites used in the model is shown in Figure 4.1. Typically, emergence of weevils begins around Julian day 200 (early June) and is completed by Julian day 300 (late September). Emergence is more advanced at some sites (e.g., Summerhill) and retarded at others (e.g., Gurtnapisha).

The graphs for Ballinagee2, Deepark2 and Oakwood2 represents emergence in a second year- this data was not used in the model, but emergence in the second year begins earlier in than in the year one. It is worth noting that some sites included in the model had more recorded observations of emergence data than others.

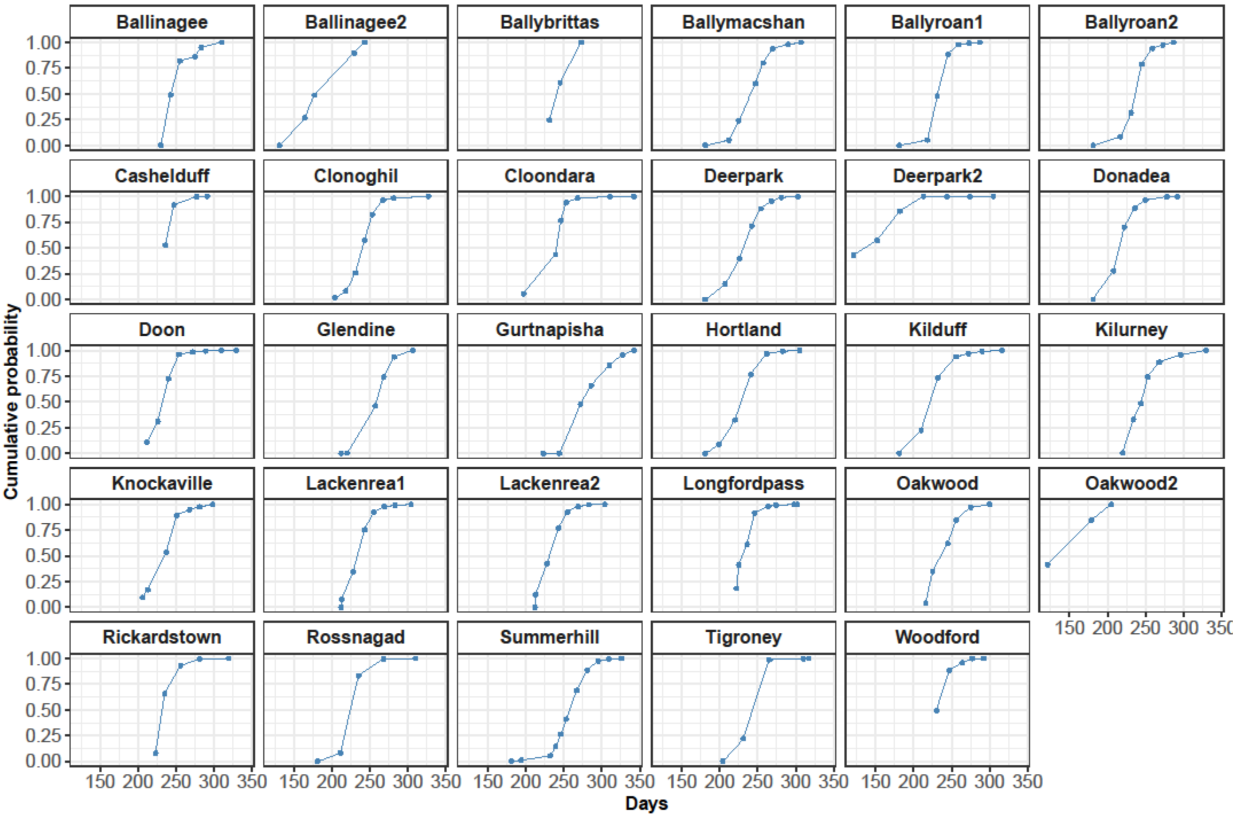


Figure 4.1: The cumulative emergence of large pine weevil on various days of the year recorded on each of the trial sites (2001-2017) incorporated in the modelling of `pineR`. Included are Year 2 emergence data for three of the sites (Ballinagee2, Deerpark2 and Oakwood2) – this data was not included in the final modelling.

4.1.1.2 Population structure (Hacking data)

Figure 4.2 shows the total number of weevils in stumps prior to first emergence at each site where it was recorded via stump hacking. The three sites with the highest populations were pine species sites. Information obtained by stump hacking is a key component of current IPM to aid foresters deciding to treat clearfell sites or not with chemical intervention (Teagasc, 2020; Dillon *et al.* 2006).

The population structure data on-site displays each development stage of larvae, pupae and adult counted just prior to first emergence. Figure 4.3 shows the proportions of weevils at each developmental stage for all sites where such data was recorded. This was useful for discussion purposes between all sites that had emergence observations (utilized in modelling) and prior stump hacking data recorded in the same year.

25 of the 27 emergence trial sites included in the forecast model had both sets of data. At most sites, a majority of the weevils were in the pupal stage (Fig. 4.3). Sites where the larval stage was dominant tended to be those at higher altitudes (e.g., Deerpark, at 260m and Gurtnapisha at 466m), perhaps indicating delayed development at those sites. However, the stump hacking was not carried out with the same data, making between-site comparisons less reliable.

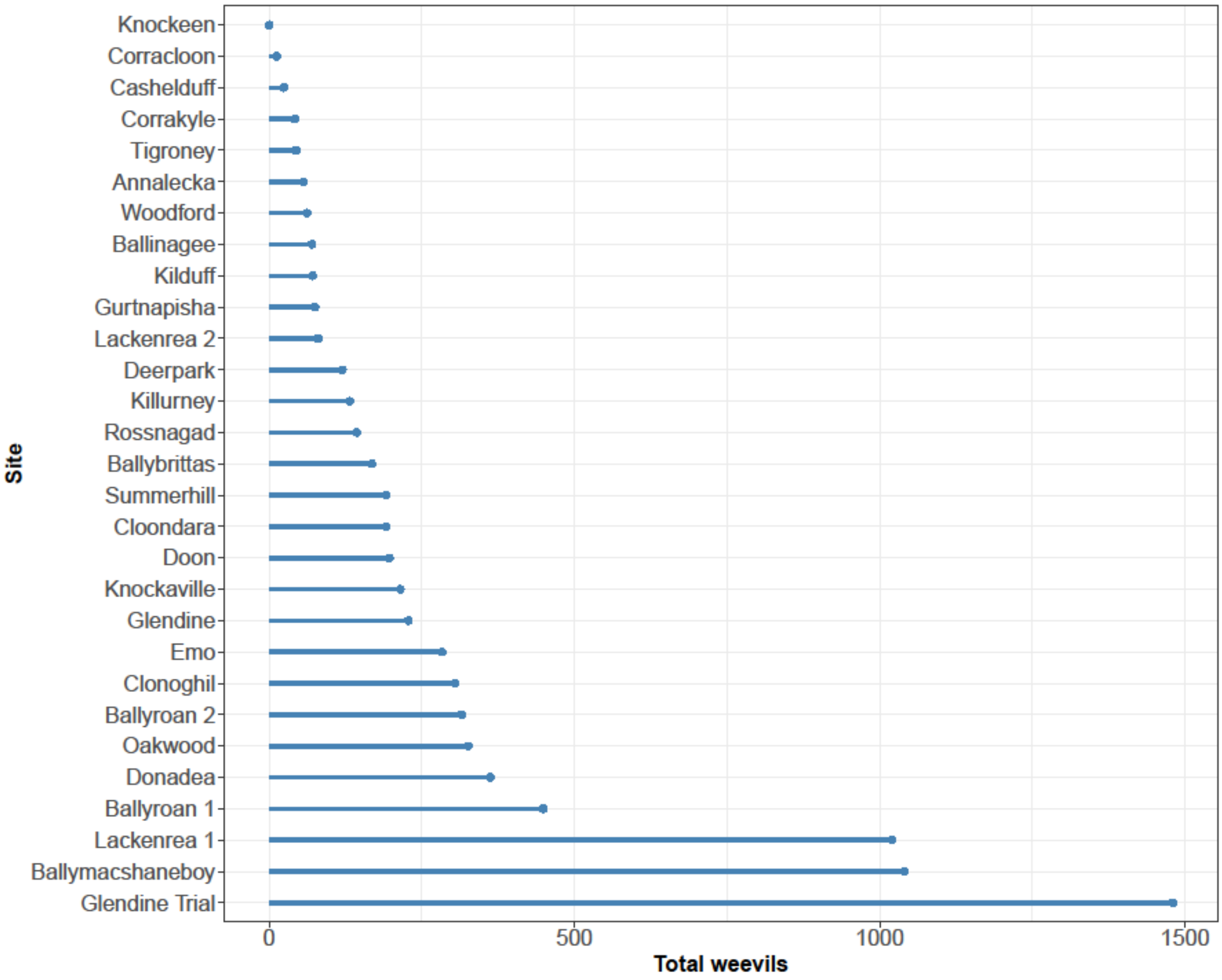


Figure 4.2: Total size of large pine weevil populations recorded at stump hacking prior to emergence on Irish trial sites. Some of these sites had no emergence data and therefore could not be included in the modelling process i.e., Emo and Annalecka.

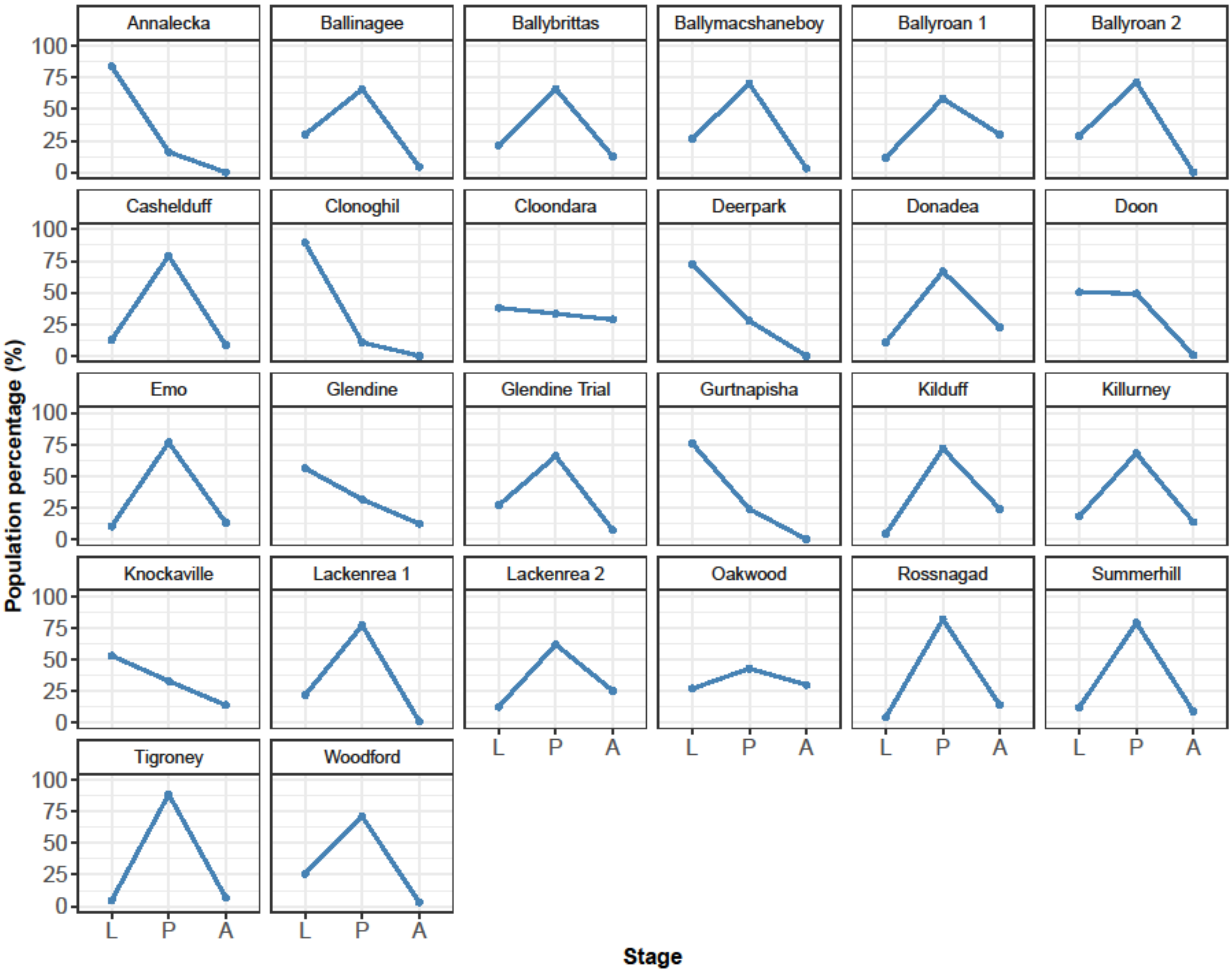


Figure 4.3: The breakdown of larvae (L), pupae (P), and adult (A) weevil stages recorded at trial sites where stump hacking was undertaking prior to emergence at Irish trial sites. Some of these sites (Annalecka and Emo) were not included in the modelling as there was no data on emergence.

4.1.2 Evaluation of meteorological data

The temperature data, obtained from the meteorological stations and used as a proxy for site temperature, was initially compared to the Met Eireann gridded datasets (0.01-degree resolution). A comparison between the meteorological station and gridded temperature data for all 27 emergence trial sites are shown in Figure 4.4. The blue line corresponds to the temperatures in the gridded climate data from Met Eireann at 0.01 (1km) degree resolution (Walsh, 2012) while the red line corresponds to the temperature data obtained from weather stations within a certain radius of the large pine weevil emergence trial sites in accordance with the rules and implementation of `pineR`. As emergence data was available for different time periods and years, the temperature comparison was undertaken for equivalent periods at each site. The time series corresponds to three years to cover the emergence and life cycle period.

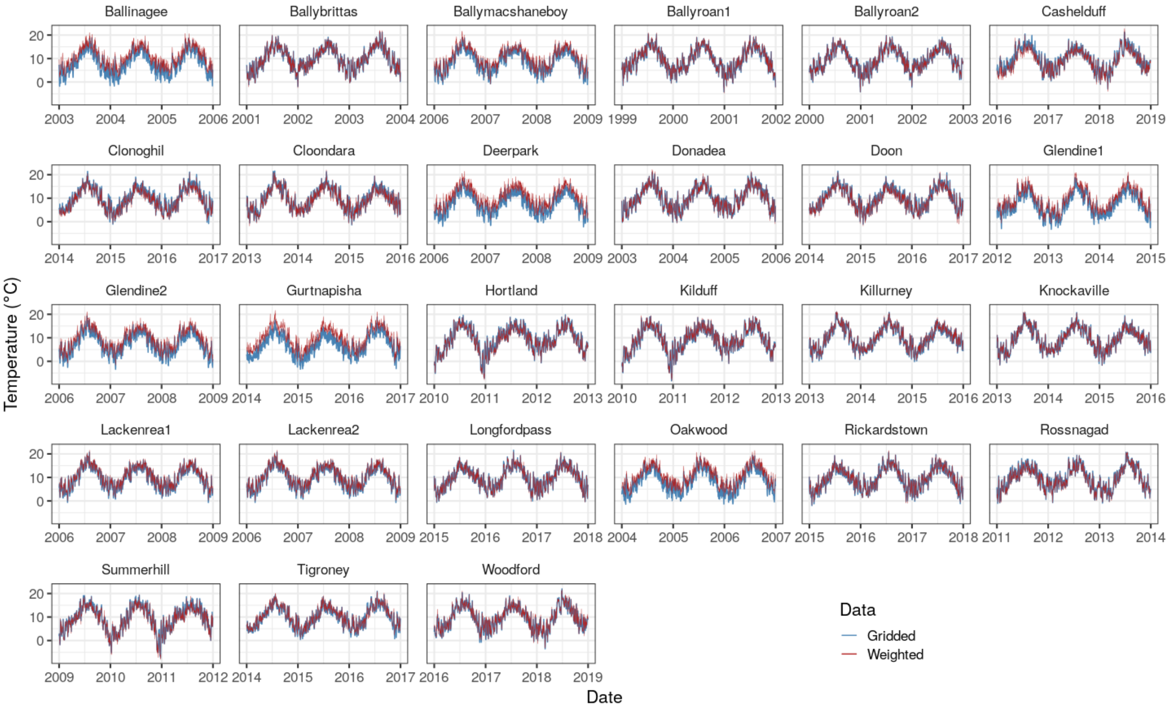


Figure 4.4: Met station temperature data at each emergence trial site compared against the gridded temperature data from the Met Eireann Grids (1km resolution) at each site taken for specific emergence trial year period on site. The facets highlighted in red represent weather station data. The blue is the gridded data (Walsh, 2012)

4.1.2.1 Comparison of met station data against gridded data.

The observed values from the nearest meteorological stations are shown at most sites to match the daily variations and patterns in the gridded data (Fig. 4.4). Figures 4.5 and 4.6 show the weather station and gridded temperatures for Hortland and Kilduff respectively for the emergence trial period (2010-2012). Both sites were monitored for weevil emergence by Maynooth researchers in 2011 with this data used in the forecast model.

Figure 4.5: A comparison of the weighted observed temperature data sourced from 4 weather stations (Lullymore, Derrygreenagh, Ballinla House and Dunsany) against the gridded temperature data for the trial site coordinates of **Hortland** for the trial period 2010-2012.

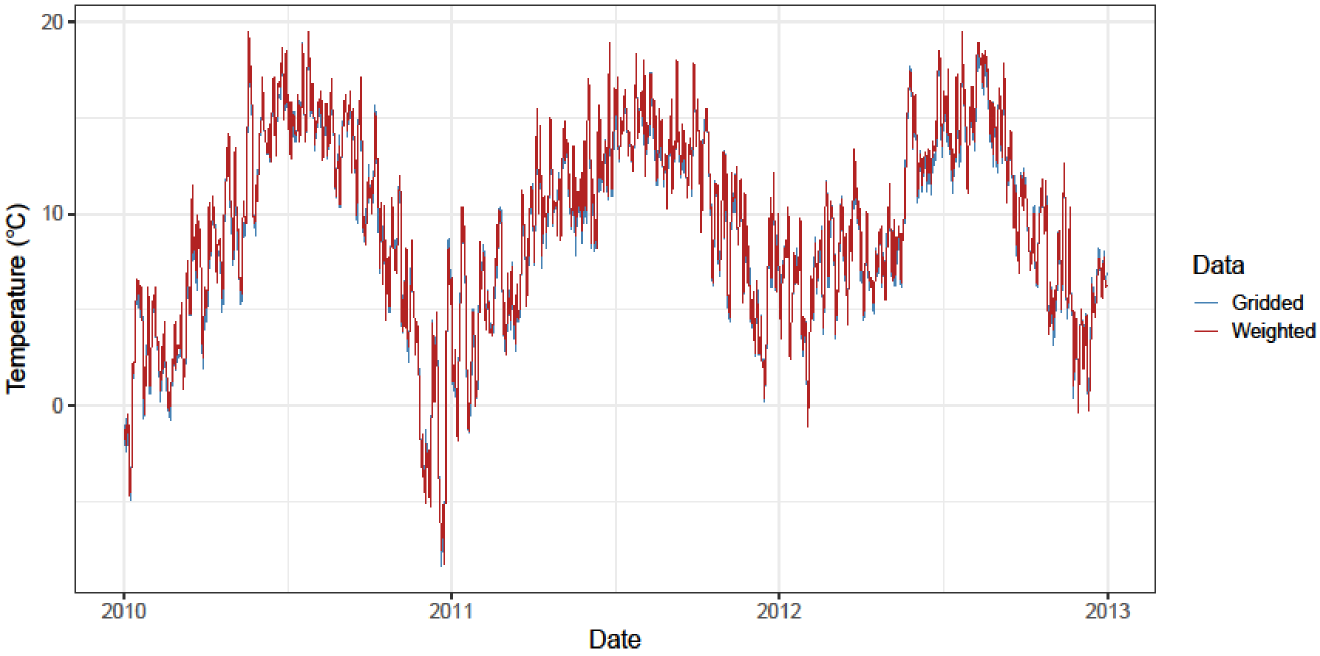


Figure 4.6: A comparison of the co-located observed temperature data sourced from 1 weather station within 10km radius as per the rules of `pineR` at Derrygreenagh (6.56 km) against the gridded temperature data for the trial site coordinates at **Kilduff** for the trial period 2010-2012.

At Hortland, the weighted temperature from 4 stations near the trial site show good agreement with the gridded weather data for the grid representing Hortland (Figure 4.5). Similarly, at Kilduff, the selected co-located weather station of Derrygreenagh, which lies within 10 km of the trial site, also shows good agreement with the selected gridded data for that location (Figure 4.6). Hortland is approximately 34.1 km from the Kilduff trial site. As these are relatively low-lying sites, the comparison between the weighted and co-located meteorological data is expected to be good.

Some sites are shown to have considerable offsets, but this can be related to site specific elements such as altitude and slope which will be discussed in Section 4.4 covering the influence of site-specific covariates. It is noteworthy that at some of the more elevated sites (e.g., Ballinagee, Deerpark, Gurtnapisha and Oakwood) values from the gridded temperature dataset are somewhat lower than those from the met station(s). Figures 4.6 (above), 4.7 and 4.8 (below) show the site comparison for gridded and met station temperature data for Kilduff (Co-located site), Deerpark (weighted) and Gurtnapisha (Co-located site). These latter two sites are interesting as they show offsets between the met station and gridded data at the higher altitude sites (Deerpark at 260m and Gurtnapisha at 466 m; Figs 4.7 and 4.8) but no notable offset at the lower altitude site of Kilduff (93m; Figure 4.6).

There are other isolated influences too in terms of the effects of imputation for missing temperature values as well as coding errors in implementing the original model at certain sites. The initial visualisation of the Met Eireann gridded data shows that temperature largely co-varies with the met station temperature values. Where there are disparities, this is investigated with a closer look at site specific elements at these locations (e.g., elevation). All this aids the process of demonstrating the utility of the 1km gridded data and establishes its use for input into the emergence forecast model. Importantly, the use of the gridded data enables the model to be transferred to sites where there is none or limited meteorological information.

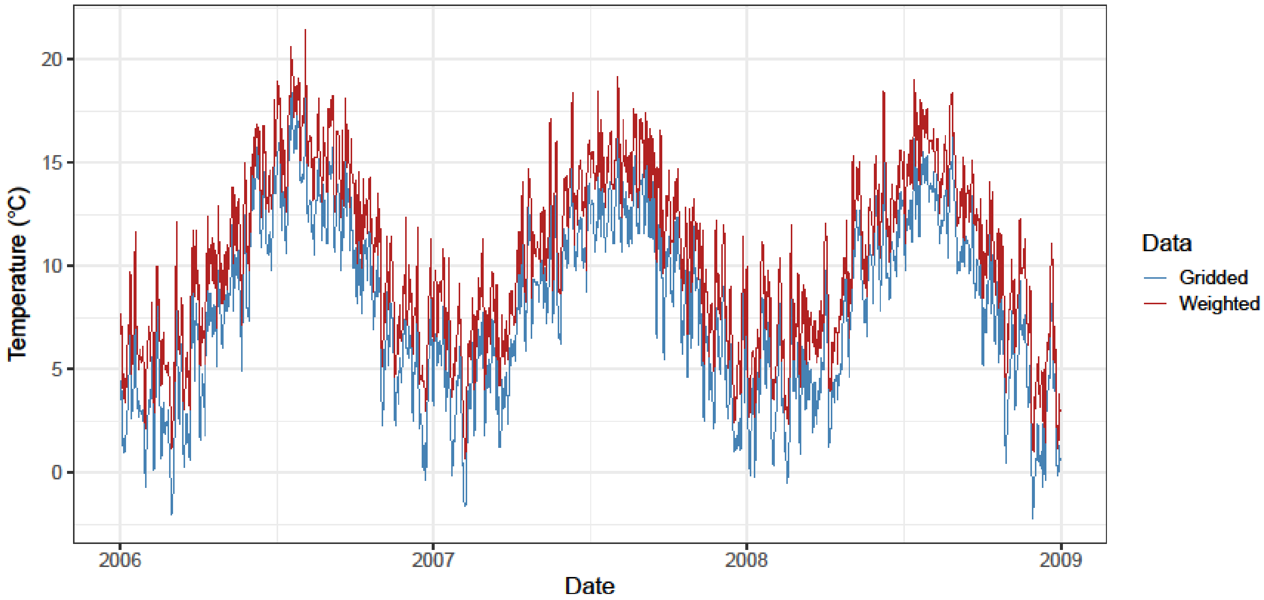


Figure 4.7: Comparison of the met station and gridded climate data temperatures for the **Deerpark** trial site (Weighted) with use of the 0.01-degree resolution (1km) data from the Met Eireann gridded climate dataset used in the final forecast modelling of `pineR`.

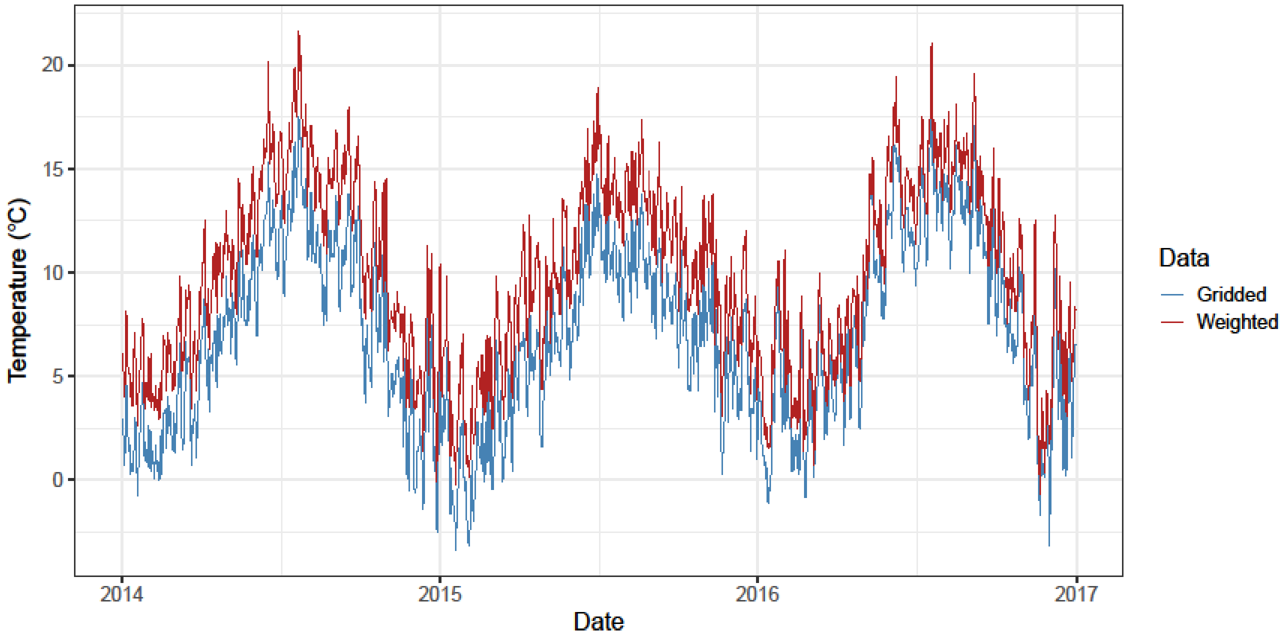
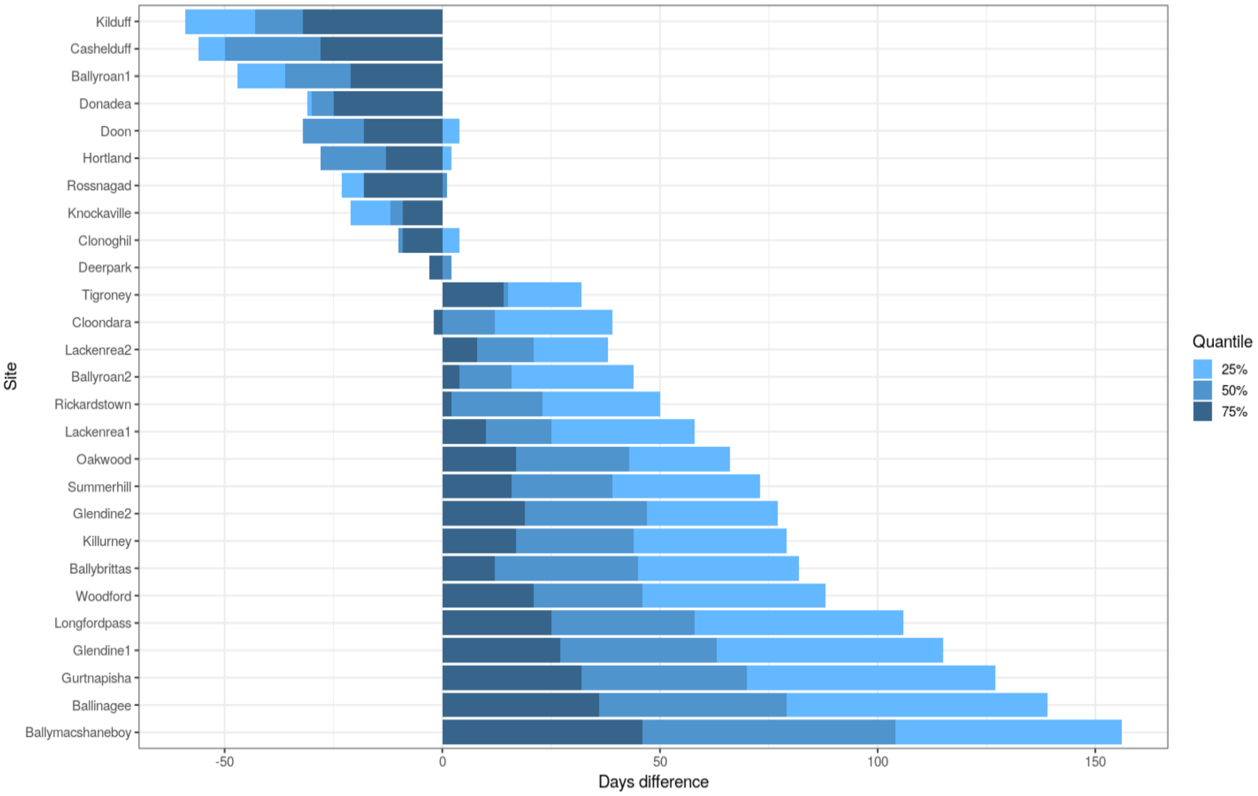


Figure 4.8: Comparison of the weighted and gridded climate data temperatures for the **Gurtnapisha** trial site (Co-located) with use of the 0.01-degree resolution (1km) data from the Met Eireann gridded climate dataset used in the final forecast modelling of `pineR`.

4.2 PineR results before correction.

The model-simulated emergence is compared to the observed emergence data to analyse how close the predicted emergence was to the emergence recorded on-site. Differences (observed & predicted) in days for the 25%, 50% and 75% quantiles of cumulative weevil emergence are shown in Figures 4.9 and 4.10 visually and figuratively in Table 4.1 for the uncorrected model. This is the model utilising both the weighted met station data and the gridded temperature data at 0.01-degree resolution (Walsh,2012) for all 27 weevil trial sites in the final forecast predictions. These values indicate a better performance of the model using gridded temperature at most sites even prior to the finalised correction. When run with the met station data the model tended to predict somewhat earlier emergence than was observed for 17 sites with later emergence predicted for 12 sites and a very close match for one site (Deerpark) (Fig. 4.9). When using the gridded temperature as input, there was a more equal distribution of early and late predictions across sites (Fig. 4.10).



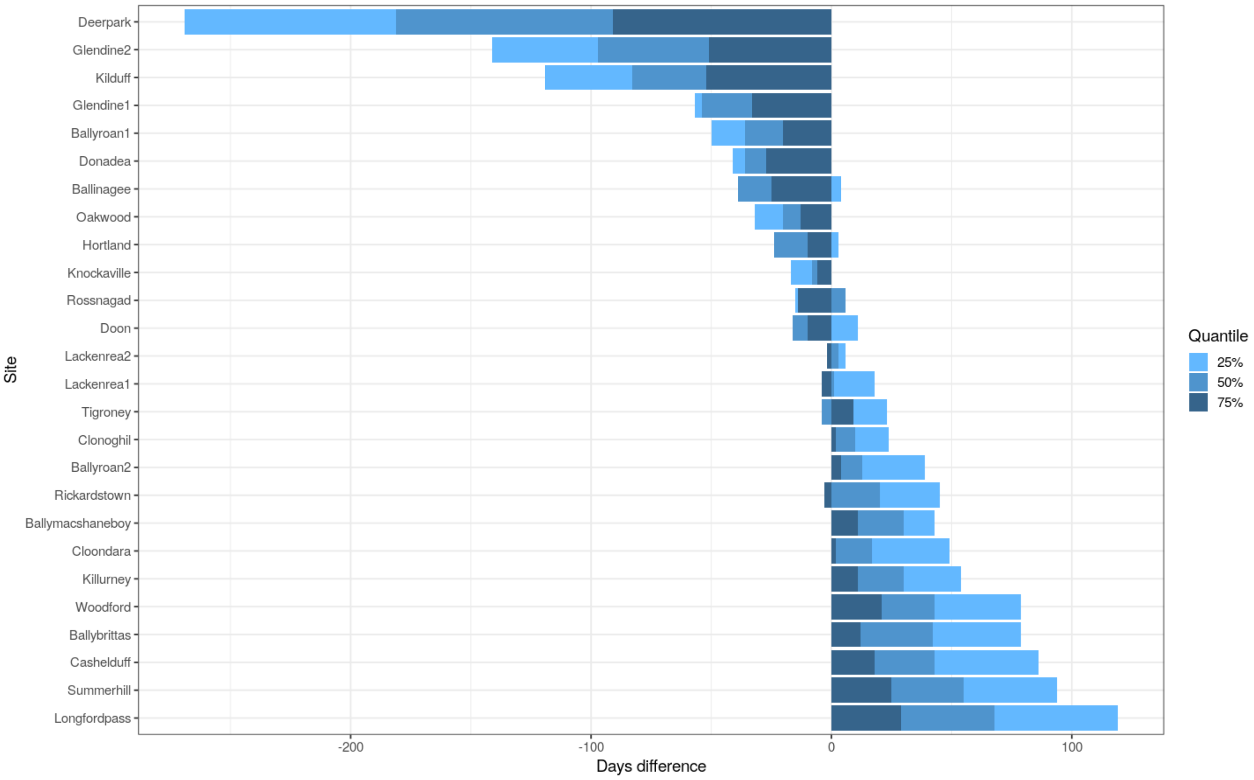


Figure 4.9 & 4.10: The difference in days at weevil trial sites in Ireland in the uncorrected model utilising weighted/co-located data (Top) and the gridded model at 0.01-degree resolution (Walsh, 2012) (Bottom). Values are Observed days – Predicted days for 25%, 50% and 75% of weevils to emerge, therefore negative values indicate that model-predicted emergence is later than observed.

Notably, when using the gridded temperature data in the model the predicted emergence at Deerpark was considerably later than observed emergence. This is one of the sites described in more detail in the following sections. Table 4.1 uses the same data shown in Figs 4.9 and 4.10 (observed - predicted days for emergence) but displayed on a site-by-site basis, allowing a comparison to be made of how the model performed when using met station versus gridded temperature inputs at each site. Notably, when using met station inputs, the model tended to forecast earlier emergence than observed for high altitude sites such as Ballinagee (425 m), Ballymacshaneboy (323 m), Glendine (380 m) and Gurtnapisha (466 m) (Table 4.1).

Table 4.1: The difference in days between observed and simulated emergence at weevil trial sites utilising the observed meteorological data and gridded data prior to correction in pineR. The difference in days (Observed – predicted) are shown for the 25%, 50% and 75% quartiles of cumulative emergence. Certain sites are described in more detail in the text.

Site	**Altitude (m)**	Weighted/co-located			Gridded		
		Q25	Q50	Q75	Q25	Q50	Q75
Ballinagee	**425**	60	43	36	4	-14	-25
Ballybrittas	**68**	37	33	12	37	30	12
Ballymacshaneboy	**323**	52	58	46	13	19	11
Ballyroan1	**100**	-11	-15	-21	-14	-16	-20
Ballyroan2	**100**	28	12	4	26	9	4
Cashelduff	**215**	-6	-22	-28	43	25	18
Clonoghil	**124**	4	-1	-9	14	8	2
Cloondara	**41**	27	12	-2	32	15	2
Deerpark	**260**	0	2	-3	-88	-90	-91
Donadea	**64**	-1	-5	-25	-5	-9	-27
Doon	**71**	4	-14	-18	11	-6	-10
Glendine1	**380**	52	36	27	-3	-21	-33
Glendine2	**380**	30	28	19	-44	-46	-51
Gurtnapisha	**466**	57	38	32	*	*	*
Hortland	**76**	2	-15	-13	3	-14	-10
Kilduff	**93**	-16	-11	-32	-36	-31	-52
Killurney	**371**	35	27	17	24	19	11
Knockaville	**95**	-9	-3	-9	-9	-2	-6
Lackenrea1	**190**	33	15	10	17	1	-4
Lackenrea2	**195**	17	13	8	3	3	-2
Longfordpass	**130**	48	33	25	51	39	29
Oakwood	**250**	23	26	17	-12	-7	-13
Rickardstown	**90**	27	21	2	25	20	-3
Rossnagad	**94**	-5	1	-18	-1	6	-14
Summerhill	**90**	34	23	16	39	30	25
Tigroney	**246**	17	1	14	14	-4	9
Woodford	**192**	42	25	21	36	22	21

4.2 `pineR` results before correction for selected sites.

To further examine the capabilities of the `pineR` package, certain sites were selected for analysis as being of particular interest. The selected sites were Hortland, Ballinagee, Glendine, Summerhill, Deerpark, Kilduff, Lackenrea (2 sites) and Gurtnapisha. These sites were selected for their specific differences in terms of how the sourced met station temperature data were handled (co-located or weighted) and in terms of site covariates used to assess forecast performance. Kilduff, Summerhill and Gurtnapisha had co-located meteorological sites meaning their nearest weather data was sourced from a single weather station with no weighting applied. At the remaining sites, weighted temperature data was employed. Six of the trial sites selected were pine and 3 were spruce. Altogether, only 4 spruce sites were used in the model. In terms of soil type, 3 of the selected sites were peat and 6 had mineral substrates. Hortland as a pine site is useful as a comparison as it has a mineral soil substrate.

In this section a graphical representation is given for the model run with met station and gridded temperatures shown side by side for the selected sites. Differences in days between model-predicted and observed emergence at specific quartiles are noted by reference to Table 4.1. Site factors such as elevation, aspect and slope that may have influence on the forecast model are noted. Population size and structure at stump hacking (where available) are briefly described for each site.

1. **Hortland**, a pine and mineral site located in Co. Kildare. The site is at an altitude of 76 metres and is notably flat with a slope value of 0.72%. Four weather stations were utilised to give a weighted temperature for this site (Figure 4.6). These stations were Lullymore (13.2 km), Derrygreenagh (30.2 km), Ballinla House (22.7 km) and Dunsany (19.3 km). Imputation was used to fill some missing values as outlined in Chapter 3. Using the weighted temperature data, the model estimated emergence is between +2 days and -15 days with respect to the observed emergence, depending on quartile (Table 4.1). Overall, the model indicates a slightly later emergence at Hortland compared to the observed data, but the predicted values lie within the 95% confidence interval for the predicted values (Figure 4.11). When the model was run with gridded temperature data there was a slightly better fit to the observed data being closer by 1-3 days depending on quartile than when the model used met station data (Table 4.1; Figure 4.11). There was no prior stump hacking data at Hortland.

2. **Ballinagee**, a spruce and mineral site in Co. Wicklow. The site is at a high altitude of 425 metres with a slope value of 22% at the clearfell site. The temperature input at Ballinagee was weighted, being derived from two stations, located at Ashford (22.8km) and Glenealy (24.4km). Using the weighted temperature data, the model predicts much earlier emergence at this site compared to the observed emergence data, highlighting the potential influence of low-lying weather stations to estimate higher elevation temperatures. The model predicted emergence is between +60 and +36 days earlier than observed emergence, depending on quantile (Table 4.1; Figure 4.11). The model run with gridded temperature gives a much better fit to the observed data than when run with the weighted station data, predicting later rather than earlier emergence (Table 4.1; Figure 4.11). This site had a relatively low population of weevils recorded in the emergence traps (average 1.9 weevils per trap). The population structure data at stump hacking prior to emergence on the 6th of July 2004, recorded 70 weevils composed of larvae (30%), pupae (65.7%) and adults (4.3%) at 20 stumps, an average of just 3.5 weevils per stump. As a spruce site, the low population of weevils in stumps and in emergence traps is not unexpected. The late onset of emergence and low proportion of adults at stump hacking are also not unexpected for spruce.

3. **Kilduff,** a pine and peat site in Co. Westmeath. The site is at an altitude of 93 metres with a slope value of 6% which was quite high for the region which is largely flat. Kilduff employed co-located temperature data in the model with the temperature derived from the nearest station, Derrygreenagh at 6.56 km from the clearfell site. When run with met station temperature data the model indicates later emergence than was observed on site, with a difference of between -11 and -32 days between observed and simulated values depending on quartile (Table 4.1; Figure 4.11). The uncorrected model run with gridded temperature gives a worse fit to the observed data than when run with the temperature data from the co-located weather station. While both temperature sets predict later emergence than was observed, the difference is greater using gridded temperature data (Table 4.1; Figure 4.11). At stump hacking on the 20th of June 2011, 71 weevils composed of larvae (4.2%), pupae (71.8%) and adult (23.9%) were recorded at 10 stumps on site. The site had lower observed numbers of weevils emerging (average 70.4 per stump) than nearby sites including Hortland and on a less favourable mineral substrate. The observed

emergence both simulated and observed, is relatively in line with expected trends for a pine clearfell habitat.

4. **Summerhill**, a pine, and peat site in Co. Meath. The site is at an altitude of 90 metres with a slope value of 2 % which is relatively flat. Summerhill had a co-located met station with temperature derived from the nearest meteorological station, Dunsany, 8.2 km from the clearfell site. When run with the met station temperature data the model predicts earlier emergence than observed (Figure 4.11) with an offset of 16-36 days depending on quartile (Table 4.1). The model with gridded temperature as input gives a worse fit to the observed data than with the temperature data from the co-located weather station (Table 4.1, Fig. 4.11). The site is notable for the high numbers of weevils recorded in emergence traps (mean 260 weevils/stump over the course of the year). At stump hacking prior to emergence on the 24th of June 2010, 192 weevils composed of larvae (12%), pupae (79.2%) and adult (8.9%) were discovered at 10 stumps on site.

5. **Glendine 1** is one of two trials undertaken at the site in 2007. It is a pine and peat site in Co. Offaly. Nestled in the Slieve Bloom mountains, it has an altitude of 380 metres and a slope value of over 16% - a relatively steep incline. Glendine employed weighted met data in the model with the temperature from the stations nearest the trial site at Nealstown (14.6 km), Gurteen (29.1 km) and Durrow (30.8 km). When run with the weighted met station temperature data, the model indicates earlier emergence compared to observed (Figure 4.12), with an offset of between 27 and 52 days depending on quartile (Table 4.1). The observed emergence was later than simulated possibly due to the altitudinal effect. The large offset is a repeated pattern on several high-altitude sites ran in the model. The uncorrected model run with the gridded temperature gives a better fit to the observed data than when run with the temperature data, with predicted emergence later than observed in contrast to the earlier emergence predicted using met station data (Table 4.1, Fig. 4.12). At stump hacking prior on the 20th of July 2007 (later than other hacking data), 1480 weevils in 20 stumps were composed of larvae (26.6%), pupae (65.2%), and adults (7.1%) as well as a small proportion of weevils that had already emerged as indicated by empty chambers in the bark as a percentage gone (1.1%). This was the highest population recorded at stump hacking on any emergence trial site.

6. **Lackenrea** encompasses two trials of the same year at the one location with the difference being the species of tree - pine and spruce - on separate nearby plots on a mineral substrate clearfell in County Waterford. The altitudes are similar, but the spruce plot (175m) is slightly more elevated than the pine plot (170m). The slope value for the trial site coordinates was over 11%. Lackenrea employed weighted temperature data for both trials with the same temperature data from the nearest weather stations of Moorepark (30.8 km), Dungarvan (11.4 km) and Ballincurrig (31.9 km). This site is useful for exploring the effect of tree species in the model. However, trends are similar for both tree species. Observed emergence was similar in both pine and spruce, with the simulation model (same temperature source) estimating earlier emergence compared to observed for both pine and spruce (Fig. 4.12). Initially, there is a greater divergence between simulated and observed for pine than for spruce, with a divergence of 33 days for **Lackenrea1 (pine)** compared to 17 days for **Lackenrea 2 spruce** at the 25 % quartile, but at higher quartiles the divergence between simulated and observed was similar for both species (Table 4.1). For both sites the predicted emergence is a better fit to observed when using the gridded temperature data than when using the met station data (Table 4.1, Fig. 4.12). It is worth noting that for both temperature data sets there is a closer fit between observed and predicted for spruce than for pine (Table 4.1).

   At stump hacking on the 10th of July 2007 there was a marked difference at the two sites in accordance with tree species. In the pine site, 1020 weevils composed of larvae (16.4%), Pupae (57.8%), Adult (0.8%) and percentage gone (25%) were observed at 10 stumps on site. Comparing to Lackenrea 2 (Spruce), the figure was much less with only 80 weevils recorded on site. This was comprised of larvae (11.1%), pupae (55.6%), Adult (22.2%) and gone (11.1%) across 10 stumps. It is worth noting also that there was a higher population in the pine site compared to the spruce but more importantly the indication of early emergence given by the percentage of empty chambers representing 25% of the in-stump population on the pine site compared to just 11% in the spruce. Incorporating a correction in the model to take account of population structure might change the predicted emergence values for the two tree species in future (see section 5.3).

7. **Deerpark**, a spruce and mineral site in Co. Wicklow. The site is at an altitude of 260 metres with a slope value of 23.01%, the steepest slope of all the sites used in the forecast model. Deerpark employed weighted temperature in the model with the data derived from two stations at Ashford (17.2 km) and Glenealy (25.2 km). There was a very close fit between predicted and observed emergence values, with observed values always within the 95% confidence interval of the predicted (Figure 4.12). The difference in days between the observed and estimated emergence was small (-3 to +2 depending on quartile (See Table 4.1). This was the best performing site in terms of forecasting emergence with the use of weighted data within the uncorrected model. This was noteworthy as the site was elevated and steeply sloped. When the gridded temperature data is employed, the model is over 70 days off at first emergence (Table 4.1, Fig. 4.12). The predicted run indicates much later development compared to the observed data. This site was the worst performing site in the whole forecast model utilising the gridded data, while also being noteworthy as the best performing site when using temperature data from the local synoptic weather stations. At stump hacking on the 3rd of July 2007, 120 weevils composed of larvae (57.1%), pupae (21.4%), adult (0%) and percentage gone (21.4%) were recorded at 10 stumps on site. This relatively high percentage gone may reflect different emergence patterns to other sites that may have affected forecast results. Emergence was probably earlier than most spruce sites utilised in `pineR.`

8. **Gurtnapisha**, a pine and mineral site in Co. Tipperary. The site located near Slievenaman is the highest trial site in the forecast model with an altitude of 466 metres and a slope value of 22%. Gurtnapisha employed temperature data from co-located site with the temperature for the site derived from Fethard (9.63 km). The model indicates emergence much earlier than observed (Figure 4.13), with an offset of between 32 and 57 days between observed and simulated forecasts (Table 4.1). The much earlier development indicated by the model compared to the observed data may show the potential influence of altitude. When using the gridded temperature data, the forecast model did not run as the temperature input to the model was effectively too cold at the site coordinates or grid of interest. The temperatures sourced for interpolation at that location (the highest site within the model) did not generate enough degree day accumulation to reach the required thresholds for development due to insects optimising their micro-climate and finding suitable climate space to develop. However, if all values were raised by 1 degree for the years of trial within the sourced gridded data, the model worked effectively showing the cold sensitivity of the forecast model. At stump hacking on the 7th of July 2015, 75 weevils composed of larvae (76%) and pupae (24%) were observed at 6 stumps on site indicating the population structure on site.

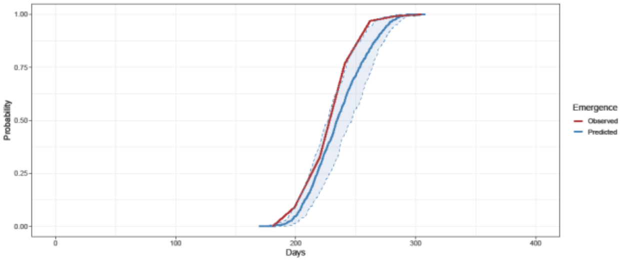
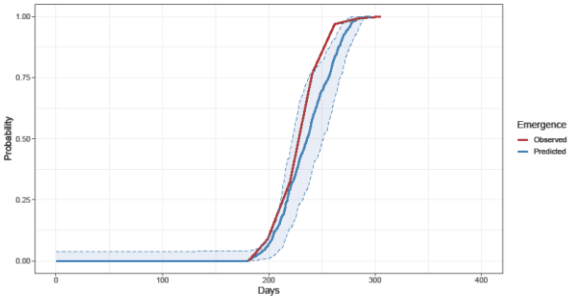
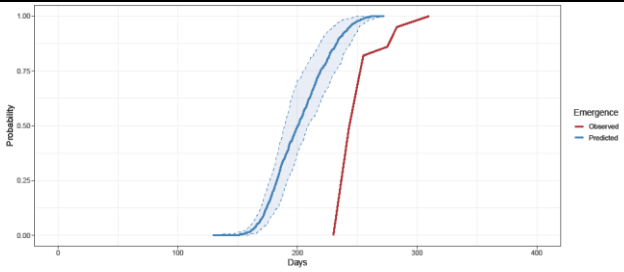
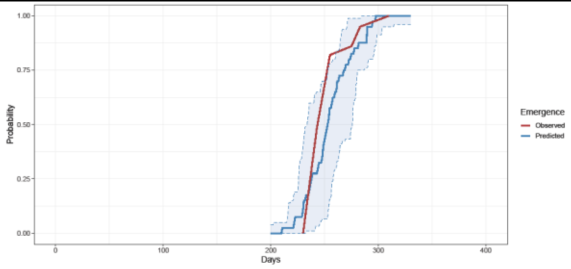
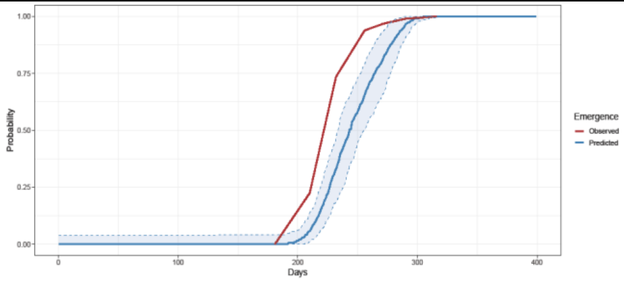
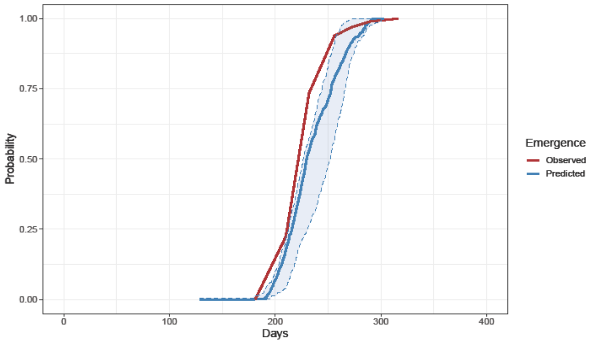
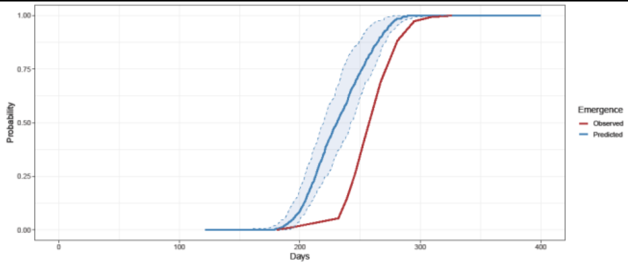
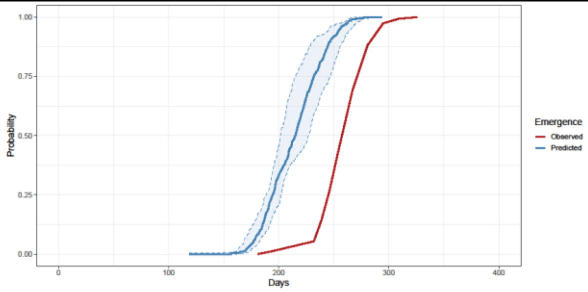
Met station temperatures as input	**Gridded temperatures as input**
Hortland	
	
Ballinagee	
	
Kilduff	
	
Summerhill	
	

Figure 4.11: The observed and predicted simulation data (with 95% confidence intervals) for cumulative emergence of pine weevil, obtained with the uncorrected model employing temperature data either from met stations (left) or the national grid (right) for four selected sites. Results for all sites are in Figures 4.14 and 4.15 respectively.

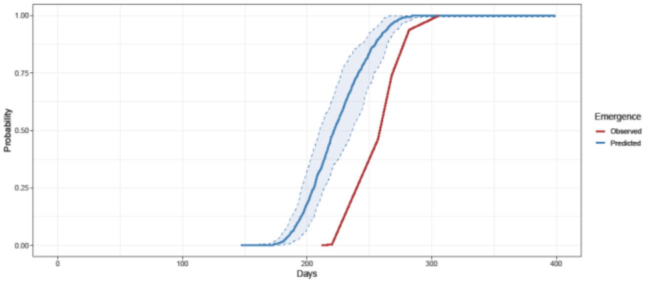
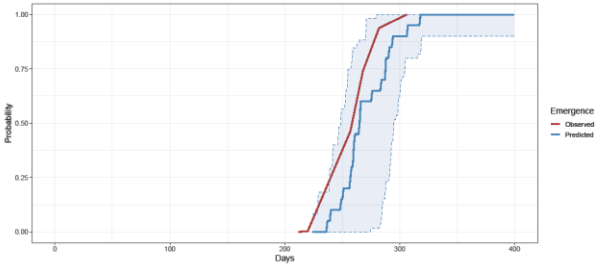
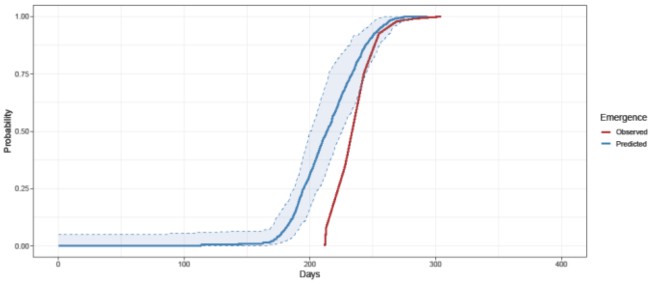
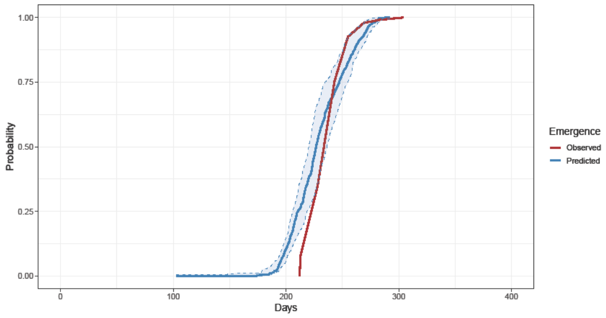
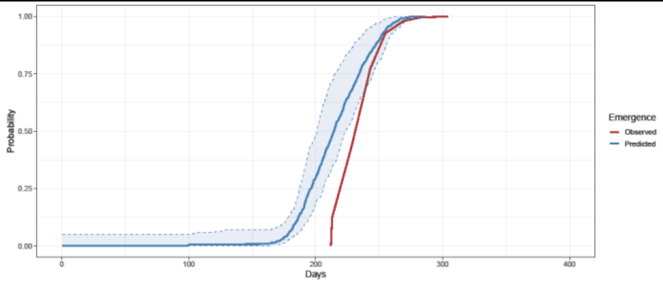
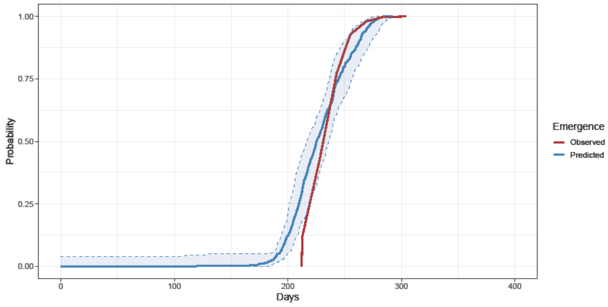
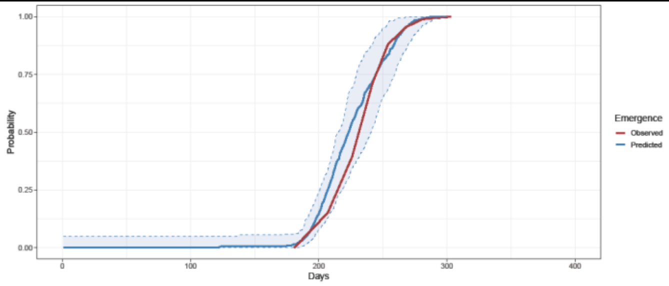
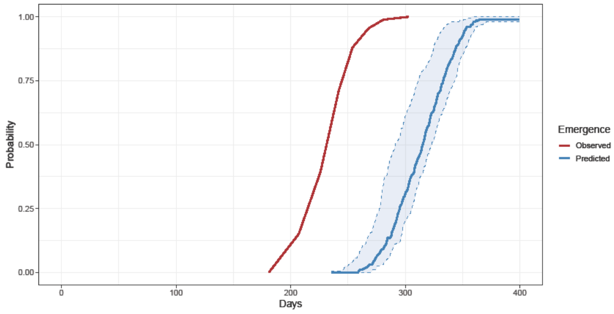
Met station temperatures as input	**Gridded temperatures as input**
Glendine 1	
	
Lackenrea1	
	
Lackenrea 2	
	
Deerpark	
	

Figure 4.12: The observed and predicted simulation data (with 95% confidence intervals) for cumulative emergence of pine weevil, obtained with the uncorrected model employing temperature data either from met stations (left) or the national grid (right) for four selected sites. Results for all sites are in Figures 4.14 and 4.15 respectively.

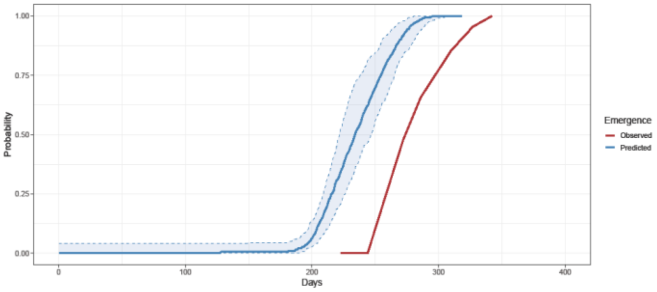
Met station temperatures as input	**Gridded temperatures as input**
Gortnapisha	
	Model did not run

Figure 4.13: The observed and predicted simulation data (with 95% confidence intervals) for cumulative emergence of pine weevil, obtained with the uncorrected model employing temperature data either from met stations (left) or the national grid (right) for Gortnapisha. Results for all sites are in Figures 4.14 and 4.15 respectively.

4.3 `pineR` results: Post correction

To correct the difference between the observed and the predicted number of days, a machine learning model, using a Random Forest method was employed. This approach used five site covariates (altitude, slope %, soil type, aspect, and tree species) to see whether they could provide useful information to correct the model estimates. In examining the correction at various quantiles of the cumulative emergence data, the results at the 50% quantile generally provided the best fit of model-simulated to observed values across trial sites according to LOOCV (Leave one out cross validation). The effect of correction is assessed by comparing outputs of the model before and after correction, firstly when run with met station temperature data and then when run with gridded temperature data as input (Figures 4.14 and 4.15, respectively). In general, the simulated emergence from the corrected model was a closer fit to the observed emergence, especially when using the met station temperatures as input. The effect of the correction on model output is described below for some of the sites previously described for the uncorrected version of the model.

4.3.1 `pineR` results comparing uncorrected and corrected model using the meteorological station data at selected sites

Figure 4.14 shows the output from the corrected model together with the outputs from the uncorrected model for comparison. At some sites, such as **Hortland**, correction has little effect on the forecast emergence of large pine weevil against the observed data. At other sites, there is a slight improvement with the correction, for example at **Kilduff**, the corrected model shows an improvement on the uncorrected model estimates, but the observations generally still lie outside the confidence interval of the predicted.

At both **Lackenrea 1** (pine) and **Lackenrea 2** (spruce) the correction slightly improves the forecast of emergence against the observed data. However, at several sites there is a more marked difference between the uncorrected and corrected model output (e.g., Glendine 1). At **Gurtnapisha**, the correction greatly improves the forecast of emergence against the observed data and there is a similar marked improvement for other high-altitude sites such as Ballinagee and Ballymacshaneboy (Fig. 4.14). It is noteworthy that some sloped and elevated sites remained offset even after the correction method was applied, suggesting that some other factor maybe be important on sites with these characteristics.

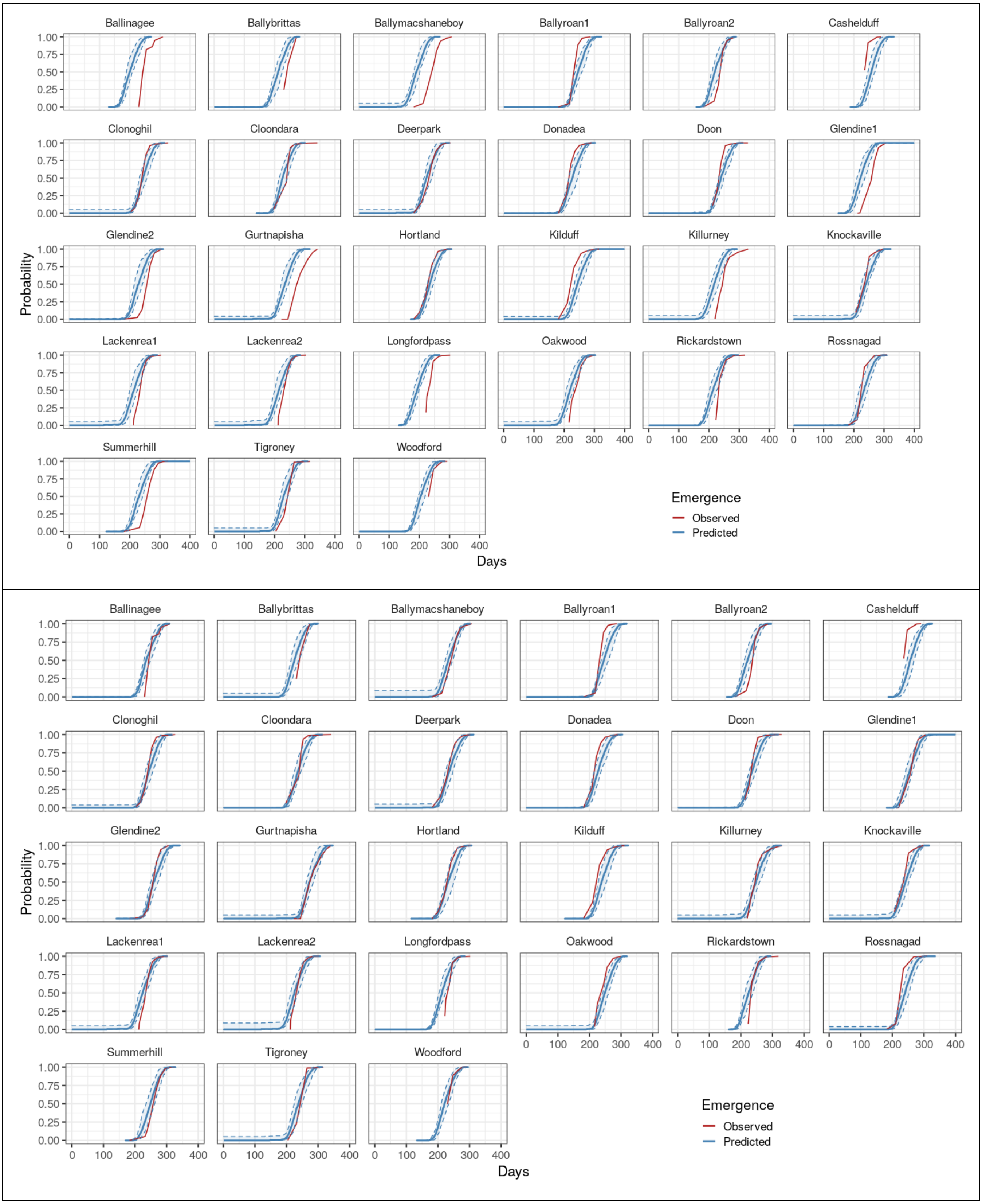


Figure 4.14: The observed and predicted simulation data (with 95% confidence intervals) for cumulative emergence of pine weevil at all trial sites, using **met station temperature data.** Above: Uncorrected model Below: correction is applied on the model.

4.3.2 `pineR` results comparing uncorrected and corrected model using the gridded temperature data at selected sites

In this section, the performance of the model before and after correction will be investigated using the Met Eireann gridded data driving the forecast predictions of the final model. The model was investigated at 25%, 50% and 75% quantiles to ascertain how much correction was needed at each site by identifying the difference in days through weevil development between observed and predicted forecasts at each trial site utilised in the modelling of `pineR`. With the final gridded interpolation data, it was again shown that in applying the correction at various quantiles of the data, the correction at 50% provided the best fit for the model across trial sites.

The outputs of the corrected model using gridded temperature data as input are shown in Figure 4.15, where they can be compared with the output of the uncorrected model also shown there. The results are shown for selected sites previously described: Hortland, Ballinagee, Kilduff, Summerhill, Glendine 1, Lackenrea and Deerpark. For Gurtnapisha, as previously mentioned, the model did not function with the gridded temperature data as input in most sites. Running the model with the gridded data seemed to improve the forecast in comparison to the weighted temperature model (derived from nearby synoptic weather stations). This is noticeable in the uncorrected runs of the model as can be seen by comparing the upper panels in Figures 4.14 and 4.15 and as previously described in Section 4.2. The applied correction further had offsets reduced (or not greatly altered) as with Glendine 1 but some sites did not improve their forecasts in terms of correcting the gap in days (differences in days to observed data) as shown at Lackenrea 1 and possibly Lackenrea 2. It is also noteworthy that Hortland and Glendine 1 for instance that the corrected model improves the initial first emergence prediction almost lining up with the observed data showing the effectiveness of the corrected forecast model at these sites. A quick overview is presented below of selected sites. Selected Individual sites with corrected and uncorrected panes are shown side by side for both the met station and gridded data are shown in the appendices for figures A3 and A4.

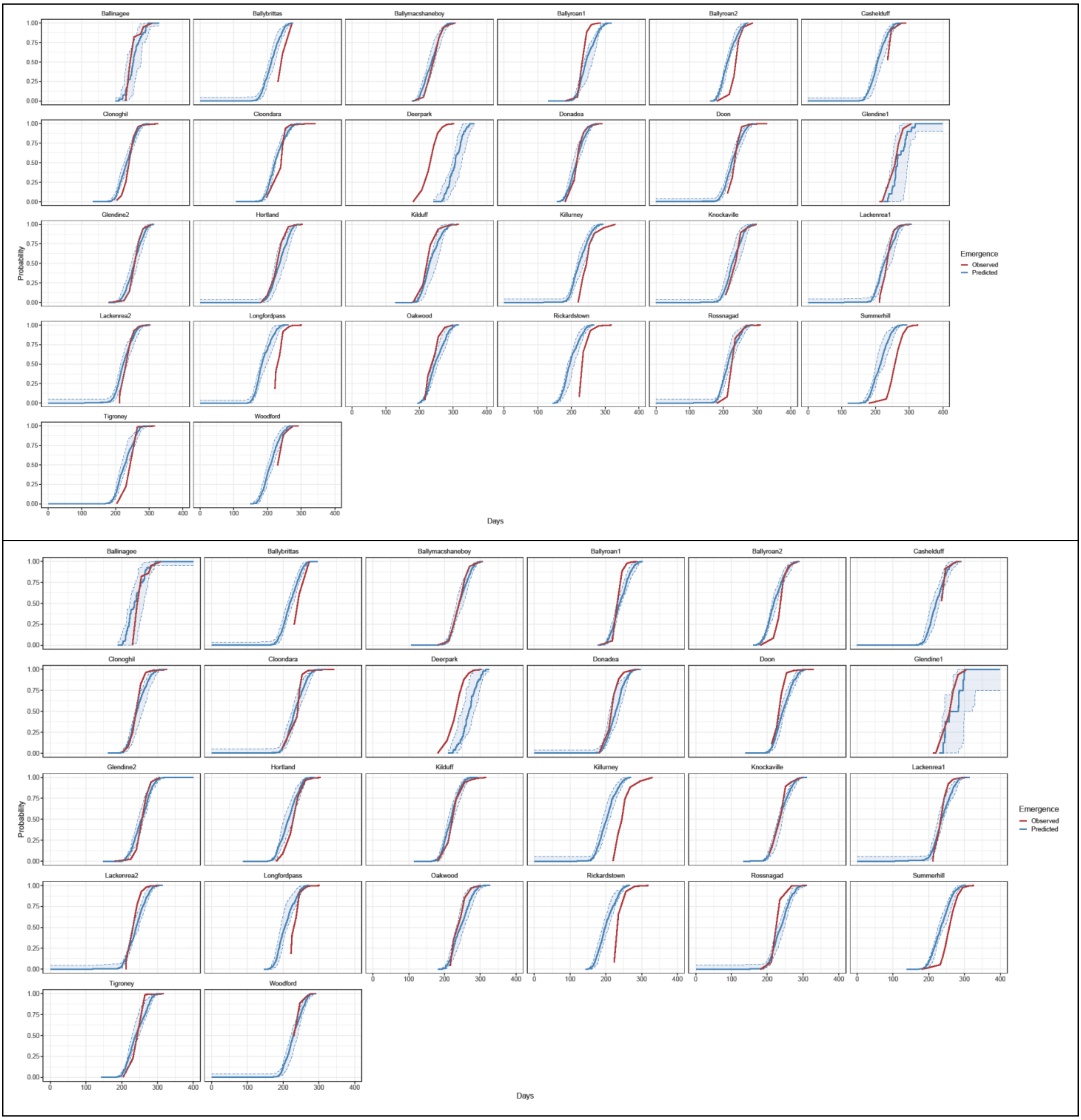


Figure 4.15: The observed and predicted simulation data (with 95% confidence intervals) for cumulative emergence of pine weevil at all trial sites, that utilised **gridded temperature data**. Above: Uncorrected model Below: correction is applied on the model.

1. **Hortland**: At Hortland, the model was close but not as smooth fitting as the uncorrected model using gridded data which was very accurate for first year emergence. The correction does not overly affect the forecast of year 1 emergence of large pine weevil on site against the observed data (Fig. 4.15).

2. **Ballinagee**: This site is a noteworthy example of the corrected model improving the uncorrected model with gridded temperature data which was a considerable improvement over the use of met station data in `pineR`. With the gridded data, the model which when uncorrected had a slightly earlier forecast for the onset of emergence is shown to be a close fit with correction applied (Fig. 4.15). However, at the 25% quartile the uncorrected model fits the observed data more closely.

3. **Kilduff:** Both forecasts at **Kilduff**, a notable improvement was recorded in the fit between predicted and observed emergence when the correction was applied to the model.

4. **Summerhill:** This site had a similar affect as per above with correction applied.

5. **Glendine 1** (2007 trial): This site was like Ballinagee, being an elevated site (and notably sloped). Although the forecast was improved with the gridded data compared to the met station data in the uncorrected iteration of `pineR,` the fit was further improved with the corrected model applied here (Fig. 4.15). The forecasting of initial emergence is in line with the observed data but there is a notable divergence later in development.

6. **Lackenrea 1** (pine) and **Lackenrea 2** (spruce): Given the proximity of the two sites, they utilised the same grid of interest from the gridded climate data. At Lackenrea 1 (Pine) the correction does not seem to greatly alter the prediction of first emergence (possibly slightly increased difference). Figure 4.15 shows the correction applied to the model did marginally improve the forecast for Lackenrea 2 (Spruce) for the early part of the emergence curve (below 50% emergence) but above that it was not improved. Using the gridded data overall did significantly improve the pineR forecast compared to the use of the met station data at this trial site as well.

7. **Deerpark**, which was the best performing site in the weighted iteration of `pineR` (Fig. 4.9, Table 4.1), was found to be the worst performing site using the gridded temperature data in the uncorrected model, with an average nearly 90-day difference between observed and predicted emergence (Fig. 4.10, Table 4.1). With correction, the forecast model improved the simulated cumulative emergence curve bringing it closer to the observed data on site, but a substantial difference remained (Fig. 4.15).

4.4 Importance of site-specific covariates in `pineR`

This section will cover the cross validation of site-specific covariates used in the forecast model to assess which features have a greater impact on the model estimates of large pine weevil emergence in terms of their own effect and by interaction with other variables. The random forest method applied assesses the difference between the observed and the predicted days of emergence. The model results for variable importance shown here are at the 50% quantile as this represented the best fit for predicted emergence. The other corrections at 25% and 75% for the model run with met station temperature inputs are included in the appendices and referred earlier in Table 4.1. One of the variables, tree species, is already incorporated from the original UK model as a function (Wainhouse *et al.*, 2014), with weevils developing faster in pine than spruce. Tree species is an important variable but not as important as the other variables in altering or affecting the differences between observed and predicted values.

Figures 4.16-4.17 show heatmaps displaying the importance of site-specific covariates in the model using met station and gridded temperature data, respectively. This was applied through the random forest method for the 50% quartile in line with the best fit for the forecast model. The diagonals reflect the importance of each variable while the off-diagonals indicate interaction between the variables. For the model runs employing the met station data (Fig. 4.16), altitude has the greatest influence on the forecast model and has been shown previously to have some effects on model performance at specific sites depending on elevation (e.g., Ballinagee, Gurtnapisha). Slope follows this, showing the effect of steep sites. Aspect, soil type and tree species are not important individually, however, the interaction between aspect and tree species appears to have influence followed by the interaction between slope and soil and then the interaction of slope and aspect in that order. The other two-way interactions are of little relevance to the results. Soil type is not important in this heat map.

Tree species has the least influence, but this is related to the fact that species is already considered in the original model (Wainhouse *et al.*, 2014) that was adapted for forecasting first emergence with `pineR`. It is also true that `pineR` predominantly is utilised here on pine trial sites. The interactions seem to be less relevant than the indicated importance in terms of affecting the forecast ability of the model for first emergence. When the model is run using gridded temperature data (Figure 4.17), a similar picture emerges, except that slope has a greater influence than altitude but both remain relevant. Tree species and aspect seem to have an impact as well but not of notable importance to the forecast model predictions.

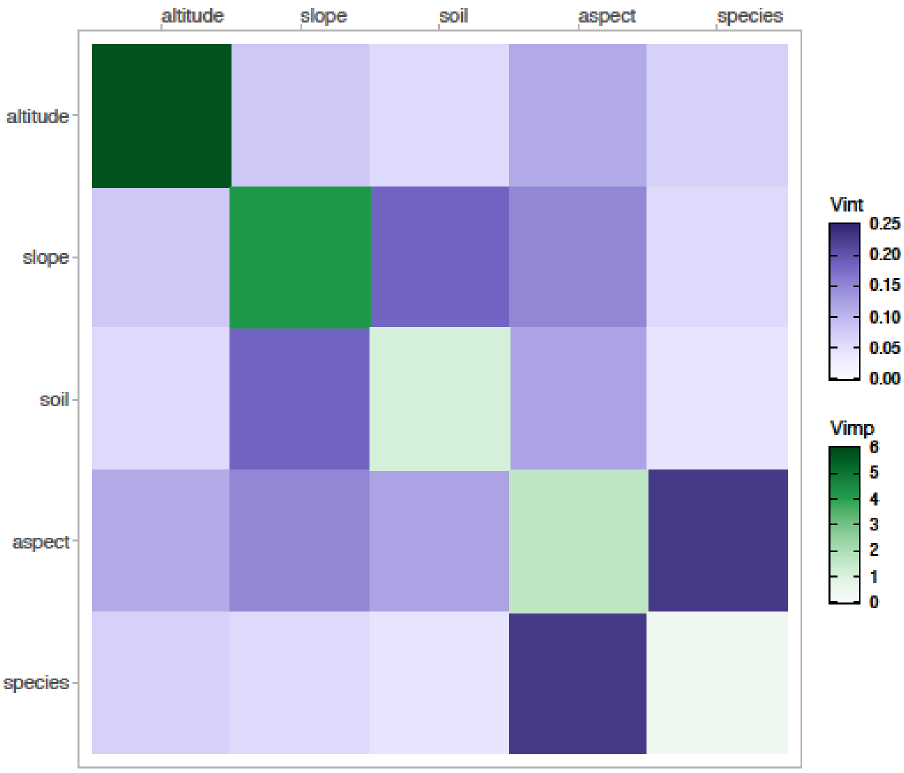


Figure 4.16: Variable importance and interaction of site-specific covariates adopted in large pine weevil (*Hylobius abietis)* final corrected emergence forecast model utilising met station temperature data. Correction here is applied at 50% which was the best fit for predicted emergence.

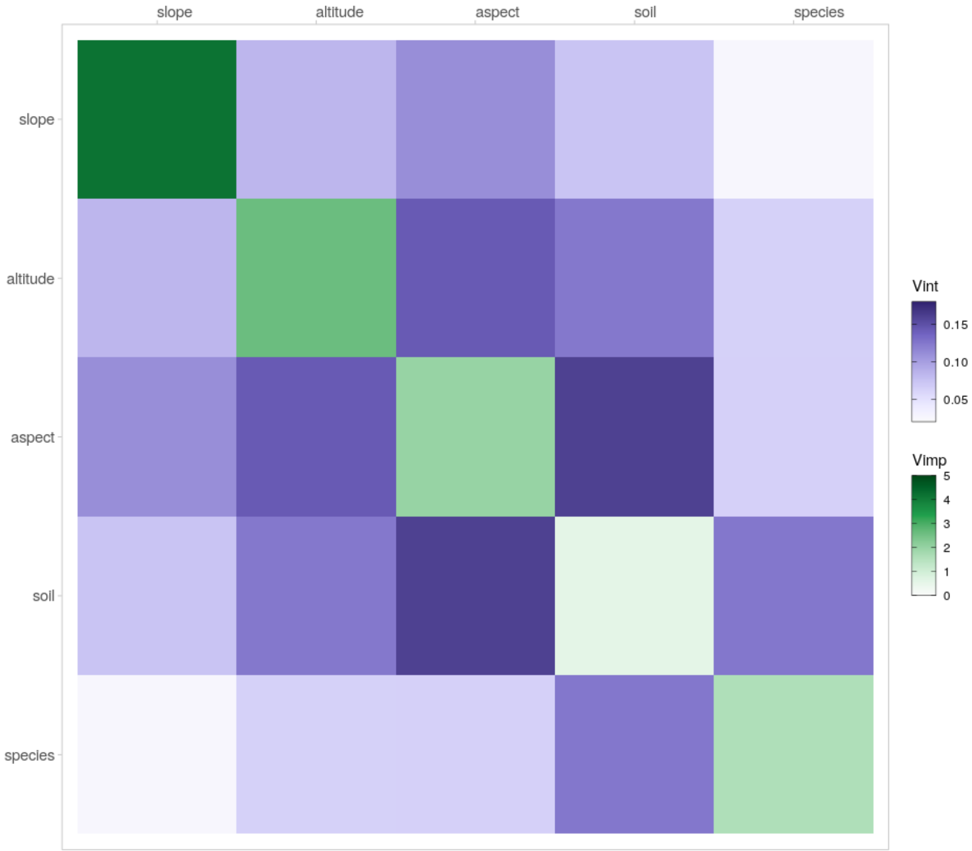


Figure 4.17: Variable importance and interaction of site-specific covariates adopted in large pine weevil (*Hylobius abietis)* final corrected emergence forecast model utilising gridded data (Walsh, 2012). Correction here is applied at 50% which was the best fit for predicted emergence.

5. Discussion and Conclusion

In this chapter, the core findings of the research will be discussed. Firstly, the performance of the corrected `pineR` model and its effectiveness in forecasting large pine weevil year one emergence at the 27 sites used in the model will be discussed. Limitations of the model and comparisons to previous modelling will also be addressed. Secondly, site specific covariates will be discussed focusing on specific covariates of influence on forecast simulations within `pineR`. Finally, the potential contributions of this work to integrated pest management of large pine weevil, pest forecasting and policy application from future work in the field will be discussed.

5.1 Evaluation of forecast model (`pineR`) performance

Evaluating the performance of `pineR` is needed to identify the capabilities of the model incorporating correction to estimate the year one emergence of large pine weevil in Ireland. This involves reviewing the relationship between local climate data and corresponding weevil development when the model was run using met station and gridded temperature inputs. Furthermore, site specific covariates of influence were identified that showed important differences between the iterations of `pineR` using met station and gridded temperature data that aids the evaluation of the corrected model. In the initial run of the model, the temperature values from nearest met stations (co-located or weighted) to the trial sites were used as input and compared with observed biological data (empirically recorded emergence of adult weevils). This provided evaluation of observed and forecast simulated emergence data. Following this the model was run using gridded temperature data. In general, the gridded temperature data was found to result in improved emergence forecasts, even prior to correction of the model as shown in Figs. 4.9-4.10 and Table 4.1. Examination of the results on a site-by-site basis allowed an evaluation of the model and to highlight potential differences between sites, with respect to the influence of certain site factors (covariates). Overall, when correction was applied, the forecast model was largely found to perform better across the selected sites with the Random Forest method using the 50% quantile for correction. It correctly estimated initial emergence at many sites and closed the difference in days between observed and predicted at many sites seen with the forecast model pre-correction.

There were some notable differences identified in the forecast model, most prominently seen in the uncorrected model runs. The temperature data for the lower Hortland site for example in both the weighted Met station and Gridded data followed a similar pattern but the weighted data (from lower weather stations than on site) seemed to underestimate the grid data at higher or more sloped sites such as Gurtnapisha and Deerpark. These sites contrasted with other sites but showed how the model when using the gridded temperature could change due to the on-site characteristics leading to changes in the interpolation. It is worth noting that interpolation can only be as good as the inputs into the forecast model. Therefore, in the case of mismatches, it can be because the site features are very different from what the model interprets, they are. The values generated are based on the interpolation of daily values across the network of sites. Based on this, one can calculate monthly, seasonal, and annual values. Therefore, if the underlying elevation model in built (DEM) does not correctly estimate the actual site elevation at a trial site for instance, then errors are likely due to the input elevation and not the interpolation of the surface which is based on the more regular daily inputs. A very notable trend was that many of the higher and more sloped sites (e.g., Glendine and Ballinagee) displayed offsets in the forecast modelling. It could be theorized that the 1km elevation used in the gridded temp values (Walsh, 2012; Flanagan *et al.* 2019) is having an effect at more notable high and steep sloped locations not appropriately captured in the 1km resolution DEM. This will be covered more in Section 5.2 on covariates of influence. In theory it was thought that the gridded data should perform better in the forecast model because it has several in-built components including for instance the elevation at trial site coordinates factored as a covariate in its 1km digital elevation model. This would include a functionality to drive cooler temperatures where appropriate to do so but there are potential limits.

It is true though also that the coarse gridded elevation model that is driving certain temperatures can also underestimate the actual temperature at certain locations as potentially shown in some of the weighted sites above (See Figure 4.4). For example, in the case of Glendine, model temperatures may be too warm for that site in terms of forecast performance. This could be due to the DEM envisaging the site is at a lower location. It is worth noting that as we assess the performance of forecasting at sites, we can also identify systematic bias that may over or underpredict relative differences in the forecast model. This brings about another question of the sourcing of gridded data with some sites when we look at sites that performed well at the weighted iteration of `pineR` vs with the gridded data.

One site of note was Deerpark where the model using gridded data forecast a much later emergence (nearly 90 days) than what was recorded on site at other elevated sites such as Ballinagee and Glendine1. The observed data seems to indicate a further initial emergence and development than expected at the trial site relating to the gridded data performance. This was prior to correction. When the correction was applied, the model performed more accurately at these sites bar minor limitations which will be discussed in the appropriate section below.

5.1.1 Comparison to previous modelling in Ireland and the UK

The `pineR` model was developed based on development times of different life cycle stages in a modularized approach to predict year one emergence of large pine weevil. Estimates obtained from the existing model were compared against observed emergence data. `pineR` expands on previous unpublished modelling research in Ireland using an existing life cycle simulation model (Wainhouse *et al*., 2014) produced by Forest Research UK, and applies a correction to the model using site specific factors or covariates. The relative importance of these covariates on the ability of the model to accurately forecast the emergence of large pine was assessed. The original UK simulation model followed a generation of weevils through their life cycle stages to predict voltinism based on temperature in the root stump microhabitat modelled from air temperature. The life cycle in the UK currently takes 2-3 years in the north and west, while in the south and east, a 2-year cycle predominates (Wainhouse *et al*., 2014). The UK model forecasted that a 2-year cycle would become predominant in the north and west that would reduce economic impact in that region, associated with warming temperatures. In the south and east however, a 1–2-year cycle would remain meaning the period of pest management required would unlikely reduce (Wainhouse *et al*., 2014).

Previous research in Maynooth (Williams et al., unpublished) tested the UK life cycle simulation model, which was originally designed to predict geographical variation in voltinism in pine weevil under climate change, to see first could its predictions be validated by comparison with observed timing of emergence at 11 field trial sites in Ireland and secondly to see could the model in its original form be used to predict site specific temporal patterns of emergence based on data from local weather stations. In this previous modelling effort, four sites were found to have similar patterns of emergence to the results predicted by the original UK model. Some sites showed differences with statistical testing. One site showed no significant difference with the Kolmogorov-Smirnov test, and three sites were significantly different from life cycle model predictions.

There was no significant difference between predicted and observed emergence for first day of emergence and 50% cumulative emergence and there was a high correlation (P = 0.026) between predicted and observed day of 50% emergence (Williams *et al.*, unpublished). Using the original life cycle model Williams et al. also showed, under current climates, a marked latitudinal influence on current *H. abietis* voltinism in Ireland, with the proportion of three-year life cycles varying from 50% in the north to a negligible proportion in southern Ireland populations. This is relevant to climate modelling of future climates excluded from this work but relevant in terms of the future potential for `pineR`. In the Williams et al. research, this component served as a partial validation of the existing UK voltinism model (Wainhouse et al., 2014). The `pineR` model partly builds on this work, including an additional 16 pine weevil trial sites, with its correction applied through the random forest method utilizing machine learning and bias correction to factor in site covariates of influence to aid better forecasting of large pine weevil first emergence.

5.1.2 Limitations of the model

There are potential implications of such models in terms of development for wider use in the field by foresters and planters dealing directly with large pine weevil management practices. One issue within the correction model of note was cold sensitivity which can be noted to affect phenological pest models (Hallman and Denlinger, 1998). This aspect was relevant at two sites, Deerpark and particularly Gurtnapisha, the highest site in the model at which the gridded temperature data was not warm enough to drive the forecast model. This site had to be excluded from the gridded forecast because of this. Interestingly, if the input temperature were increased by 1^{o}C, the site did function in `pineR`. The very issue of the use of gridded data versus the met station data was also another potential limitation. Figures 4.9 and 4.10 previously displayed how the difference in days at each quantile in the uncorrected forecast model for all sites performed in both iterations of `pineR` using met station data and gridded data. This would also affect issues such as how to accurately replicate the temperatures being experienced at an elevated site and equally within the weevil microhabitat. A further issue to the forecast predictions could also be the amount of data. More environmental data or variables such as soil, or slope could help apply a better correction using more variables. This would not necessarily influence weevil development but would influence the correction. The relationship of temperature and tree species in relation to the original U.K model (Wainhouse et al., 2014) would also be another item of note in terms of the model's ability to forecast emergence of large pine weevil. There was also an issue about the best fit in terms of the random forest correction to the forecast model when considering the 50% and 75% quantiles. Some sites performed better with correction at the 75% quantile even if the overall best fit was 50% with the random forest method. This is again linked to possibly colder and higher sites and how the observed temperature there is estimated.

The current work expands upon the previous research by including more sites but 27 is still a relatively small sample. The larger the number of trials, the more precise results will be. Compiling more trial data matched to independent weather station data could have possibly helped improve the statistical knowledge being generated in the forecast model, had they been available. In terms of utilising the gridded climate data in packages such as Climate4R or baseR, computation knowledge and timing limited the chance to incorporate further investigation into expanded best fit interpolation methods for the gridded climate data. More time and research could have possibly somewhat improved certain aspects such as the visualisation of results. Further research could be done in future on an expanded number of sites, should they become available, to broaden the work being investigated and deployed in areas such as bias correction within the forecast model.

5.2 Influence of covariates in "`pineR`" (elevation, slope, etc.)

The covariates of influence within `pineR` will now be discussed. These site-specific features were vital to implementing the model correction and identifying variables of importance within the model. Elevation and its relationship to source temperature data will be looked at in detail due to the primary and significant influence it had on the forecast model. This will be broken down by the heights of different trial sites. Glendine, Ballinagee and Gurtnapisha for instance are all sites considerably above 300 metres. In contrast, Summerhill, Kilduff and Hortland are much lower. The remaining covariates of influence on the model performance investigated in `pineR` will then be discussed. These were soil type, slope, aspect, and tree species. Site specific covariates form a crucial part of the proposed final correction to the model and are an important influence on the capabilities of forecasting emergence in `pineR`.

5.2.1 Influence of elevation and sourced temperature

Altitude was identified as a factor of influence on `pineR` both when examining the actual forecasts of emergence and then as identified as a site-specific covariate of importance at the model correction stage. More elevated sites within the forecast model were shown initially to match less with the observed temperature data across model runs. For instance, there is a clear mismatch between the met station and gridded data at Glendine, which is at a height of 380m. The nearest station utilised in the initial run of the model using met station data is higher than most Met Eireann stations at 211m, but it is the wider disparity of the meteorological network that is important here. This offset repeats in the forecast simulations too with higher sites displaying mismatches between observed and simulated model runs. Altitude can help explain the correction factor not working as well at these sites. The performance is improved in some cases but once above a certain height, the same problems reoccur even with the gridded data. The 1km gridded Met Eireann data (Walsh, 2012) would have lower uncertainty associated with the temperature estimates at or below 200 metres, as most meteorological stations are below this elevation. This will increase largely because of the interpolation of the temperature data based on the 1 km digital elevation model employed. As the elevation increases, there are fewer observations to constrain the interpolated values. This means it is likely that the lack of correction in some of the higher sites (notably Glendine and the Wicklow Mountains sites of Ballinagee and Oakwood) is an error of an over [or under] estimation at these elevated sites.

Glendine for example is nearly 400 metres and most Met Eireann stations are within 200 metres above sea level. This is not even the highest site in the model. The nearest weather station at Nealstown is slightly above 200 m elevation which is a climatological station about 14.6 km from Glendine, but this is very much an outlier in the wider measurement network and on its own would not have many sites round it to draw upon in terms of interpolation. It is in fact by far the most elevated weather station employed in the forecast model. With the Met Eireann gridded data, the surface values are more reliable the closer they are to sea level. In fact, most stations are between 0 and 100 metres. So, at elevations of 200 metres and above, there is less meteorological data with fewer surrounding stations. The underlying topographical values start to influence the forecast model with the 1km digital elevation model built within the 1km gridded data (Walsh, 2012) utilised in `pineR.` The height is effectively underestimated, due to the spatial resolution effectively resulting in lower elevations in the DEM with consequent impacts on temperature at elevation being underestimated at several weevil trial sites.

The mismatches at the higher sites therefore are attributable to the driving meteorology on the ground and related to the weather data inputs. Beyond this, there is also an issue of the weevil microhabitat in terms of their microclimate in the soil for example and depths of soil which is also inbuilt into the original model (Wainhouse *et al.*, 2014) adapted for this work. Seeing this occurrence at several elevated sites means we can see a consistent error and explain it through the meteorology adopted in the modelling. It is worth noting that gridded data also repeatedly performed better than the weighted data but still tended to have errors at the higher sites. In some sites that performed particularly well using met station inputs such as Deerpark, there were some disparities but, in that case, this may be related to the sourced historical coordinate data pulled from within the grids.

The distance to weather stations was identified as a potential factor of note also in the initial run of the model using met station temperature inputs. In the case of Glendine, the three weather stations utilised in the weighting were 15 km, 29.1 km, and 30.8 km away at Nealstown, Gurteen and Durrow respectively from the forest trial site nestled in the Slieve Bloom mountains which adds another factor in terms of varying regional topography. At Hortland, a relatively low-lying site at 76 m and on very flat terrain, the model performed much better, and this was repeated on other lower sites. At Glendine the model performed better with the gridded data but simulated later emergence than was observed while with the weighted temperature data the model-simulated emergence was earlier than observed. In both cases, correction improved the emergence forecast. The three stations employed in the weighting function were at much lower elevations than Glendine. The gridded data from Met Eireann (Walsh, 2012) performed a little better primarily because it has the elevation at the trial site as a covariate in its 1km DEM (digital elevation model) built into the grids. This function helps to drive cooler temperatures at these altitudes. The coarse gridded elevation model that is driving the temperature at Glendine and at other sites results in warmer estimated temperatures compared to what is experienced at these sites. This is likely the case at other examples covered such as at Ballinagee and Gurtnapisha and helps explain the offset between predicted and observed emergence, so we can rationalise this error in the model correction. Figure 5.1 shows the difference in altitude between the sourced weather stations for weighting the data and the matching large pine weevil trial sites used in the model.

Figure 5.2 shows the spread of station distances to trial sites within the model. It is worth noting that development on the higher sites is largely later in the sites analysed likely due to cooler temperatures at higher altitudes and again as said possibly the cooling mechanism in the gridded model being overestimated compared to the real temperatures experienced on the ground for weevil populations (i.e, Glendine and Ballinagee, etc.). Figure 5.3 shows an example of the difference in elevation of the sourced temperature data at Ashford weather station for Ballinagee versus the actual site elevation at the Wicklow Mountains sites. This displays how temperatures experienced could be quite different and affecting model performance. It is worth noting in the forecast model that Ballinagee, Glendine 1, Deerpark and Gurtnapisha are sites where also the average and maximum temperatures on the gridded data were lower than the other trial sites. This meant that compared to the weighted iteration of `pineR`, these sites had the greatest difficulty in model performance and in making an accurate forecast prediction.

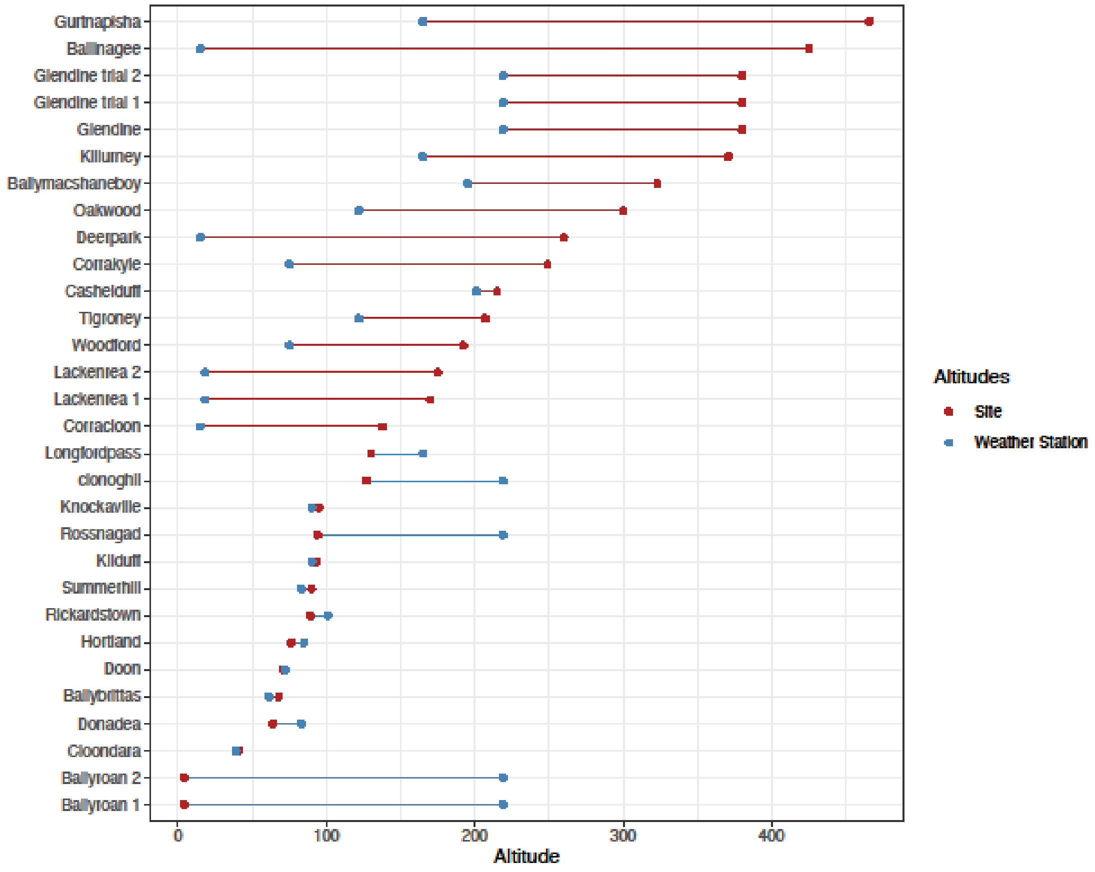


Figure 5.1: The altitude of forest trial sites and nearest weather stations utilised in the forecast model clearly showing a high disparity in height between the two at several locations.

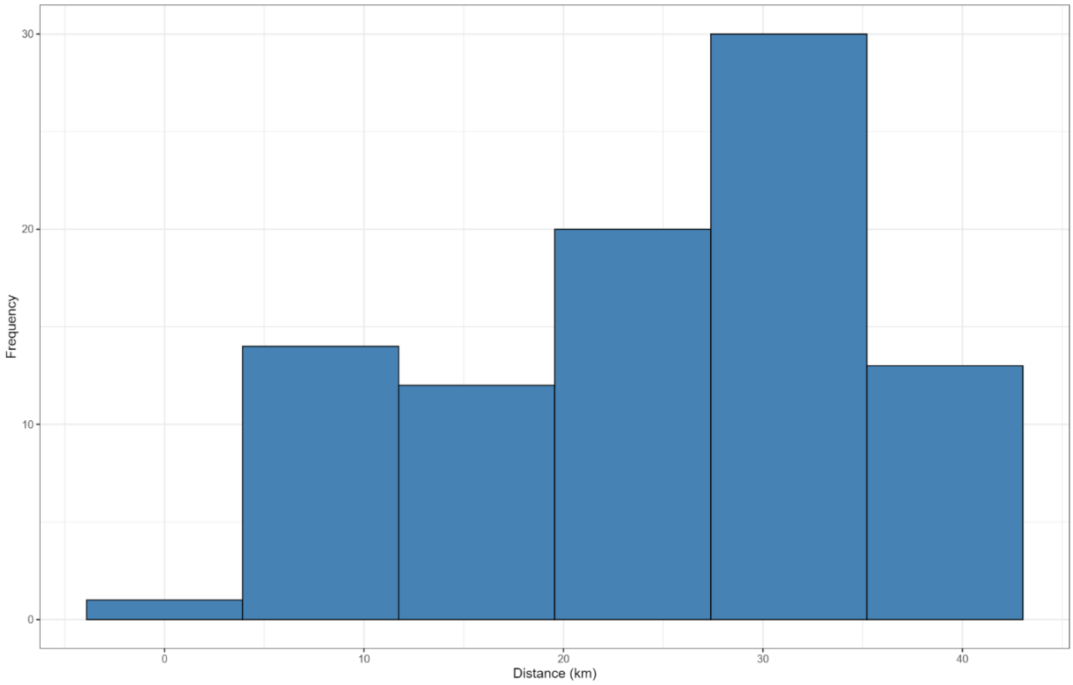


Figure 5.2: The distances of weather stations to trial sites showing the proximity in km to display where temperature data was sourced.

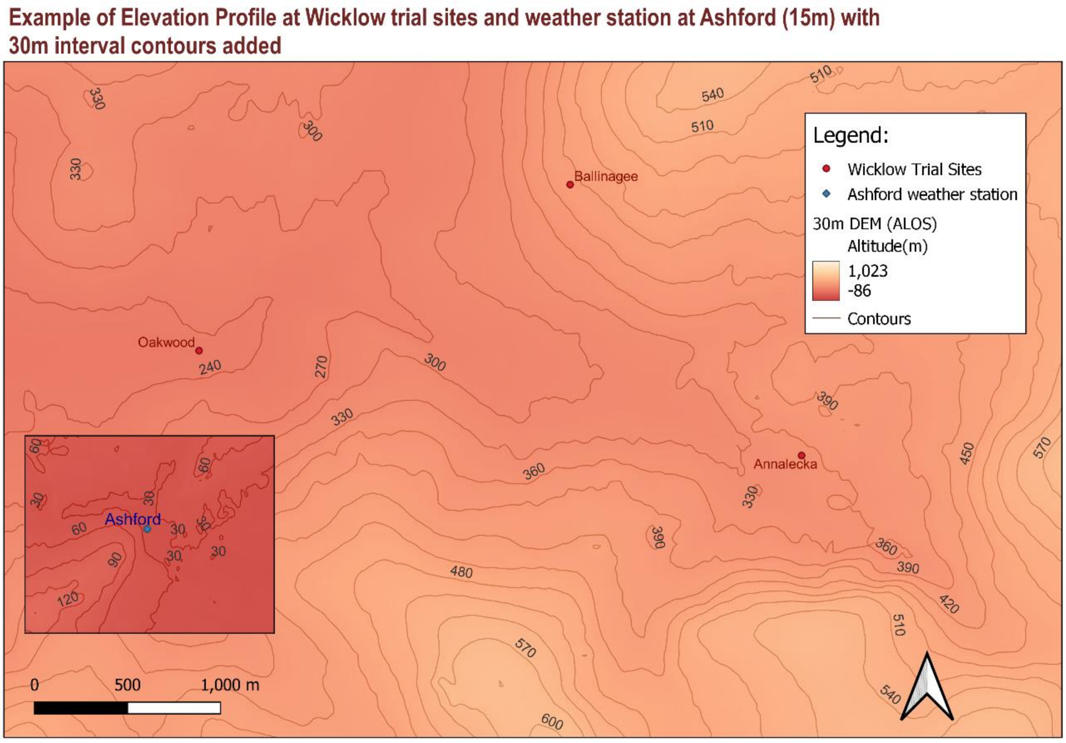


Figure 5.3: The difference of the closest weather station to the sites at Ballinagee, Oakwood and Annalecka displayed with contour lines on a DEM of Ireland (ALOS, 2020) showing the sites considerably higher than Ashford (Inset image) which is only 15m above sea level. Annalecka was not used in the model.

5.2.2 Influence of remaining covariates

The machine learning algorithm identified slope as the most influential factor when using gridded temperature data and as the second most influential factor on weevil development in `pineR` (after altitude) when using met station data. This was followed in importance by Aspect in both cases. The orientation and steepness of slopes on clear-cuts can potentially affect weevil populations due to effects on soil moisture, radiation, or wind (Tolbert, 1975). Of particular importance is the effect on solar radiation influencing the temperature experienced by the developing weevils. Slope and aspect values can potentially affect the amount of solar radiation reaching the soil (McCutchan and Fox, 1986) as well as potentially affecting the microhabitat distribution (relative to soil surface) of developing weevils depending on where the stumps are located. Figure 5.4 shows an aspect map of Ireland generated in GIS through a 30m DEM (ALOS, 2020). This was used to help aid the estimation of cardinal direction at weevil trial sites. Sites were divided by North, East, West and South facing by degrees location or nearest to as an agreed rule within `pineR`. Originally there were more variables looking at direct location according to specific aspect through GIS and DEMs that identified more accurately sites as for instance North Easterly or South Easterly. Certain flatter sites were less affected but as shown above in terms of variable importance, Aspect and Slope were relevant variables in the forecast model. North facing sites were shown in fact to be closer in terms of predicted forecasts for pine weevil first emergence in the uncorrected iteration of `pineR.` Deerpark was one site noted as an outlier in the results. It also had one of the steepest slopes as well incorporated in the forecast. In terms of its anomaly, it is worth noting that its location in a steep valley and on a northern aspect at the site may have an affect (Tolbert, 1975). A lot of headwaters of the Liffey and other water bodies are below Deerpark with notable wetlands. So, it is one possibility that the valley aspect as well as the in- built DEM in the gridded data may have its own complications in terms of affecting development rate in a model. Slope and aspect values are relevant in terms of how they may affect solar radiation reaching the soil (McCuthcan & Fox, 1986) and potentially affect the microhabitat population of weevils (relative to surface) of developing weevils depending on the location of stumps.

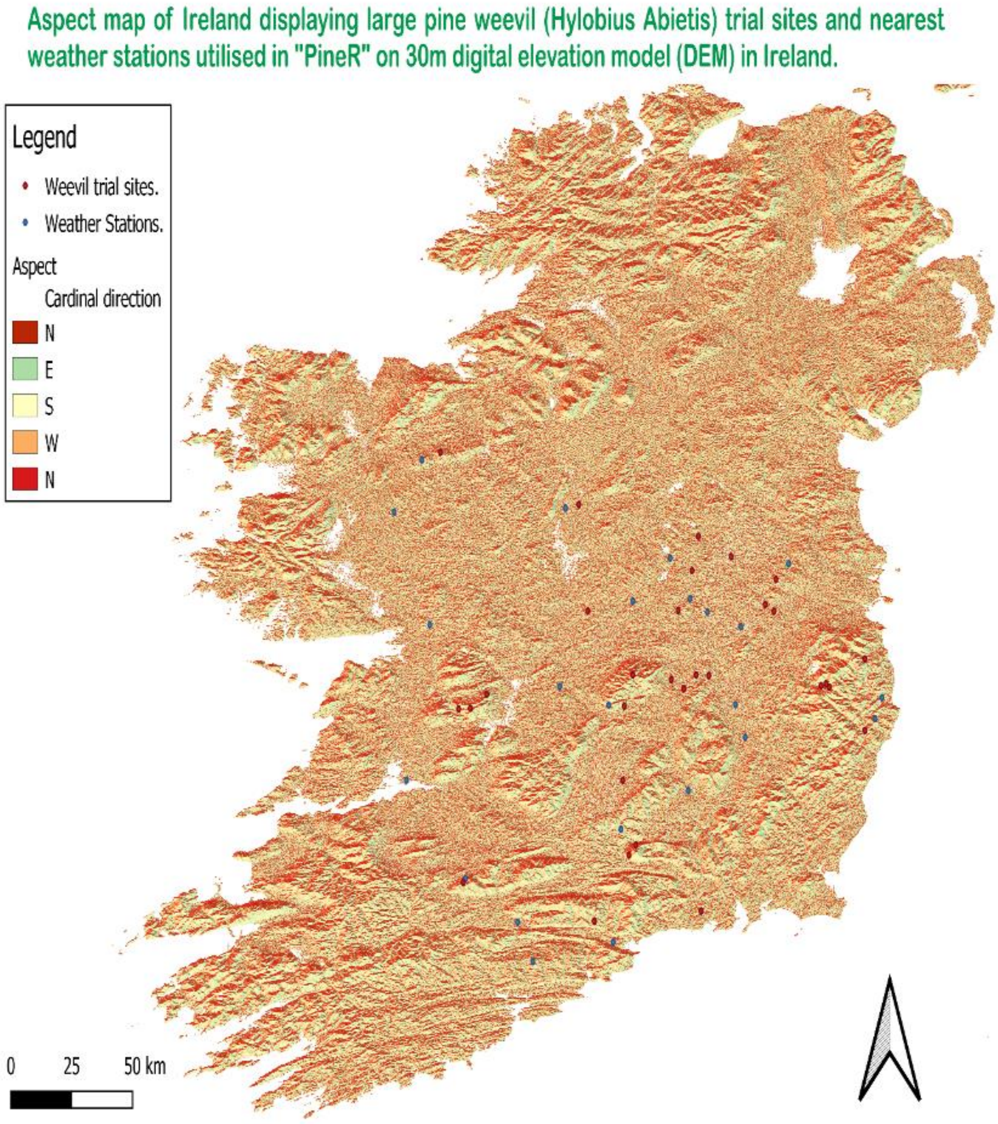


Figure 5.4. An aspect map of Ireland matching trial sites and weather stations utilised in `pineR` to Cardinal direction on a 30m DEM of Ireland (ALOS, 2020).

The large pine weevil develops faster in pine which appears to be more suitable for the species than spruce, and pine stumps also support high weevil populations (Griffin and Dillon, 2008; Kapranas *et al..*, 2017) as displayed at Glendine for instance, the highest observed weevil population site utilised in `pineR`. The results from `pineR` forecasting reflects faster development occurring in pine stumps. For example, emergence occurs quicker at Kilduff, Summerhill and Glendine (all pine sites) compared to later development at Ballinagee and Oakwood (spruce sites). Figure 5.5 shows larvae, pupae and adult weevils found in an untreated stump at one of two Coillte trial sites visited as part of background research for this thesis.

This was a pine and peat clear-fell site in County Kerry where many weevils of different developmental stages were discovered on the site visit in August 2021 compared to a spruce and mineral site nearby at Direenauling, Co. Kerry where only one adult weevil was recorded at stump hacking on the same day.

Figure 5.5: Larvae, pupae, and adult weevil that were found on some untreated plots on a Coillte trial site of a pine and peat clear-fell near Cummeenvrick, Co. Kerry on the 18th of August 2021.

As stated, stumps of most conifer species support weevil development, but the numbers developing in and emerging from pine (*Pinus* species) are far higher than from spruce (*Picea* species) in Europe and in Ireland with up to 100,000 emerging on a single site (Dillon and Griffin, 2008). Examples too of the initial on-site investigation and indeed emergence performance in the simulation model display the difference between tree species at Ballinagee and Deerpark (spruce) versus Glendine and Summerhill (pine) for instance. In the initial modelling with weighted data and somewhat overall we saw some of the spruce sites were better performing than the pine sites. In terms of variable importance, soil type and tree species were not shown to carry as much significance as other covariates in the random forest correction method applied. Soil depth was not factored into the covariate investigation as it was predetermined by the original UK model (Wainhouse *et al.*, 2014) with pine set to deeper development at 30cm and spruce 20cm as per the model methods (See chapter 3).

As previously stated, tree species is also inbuilt in the original model as a function which with most of the sites being pine could have affected the influence of “Species” on forecasting first emergence.

Soil type can influence the extent of pine weevil damage (Luoranen *et al.*, 2017; Wallertz *et al.,* 2018) but has not been reported to affect the rate of development of the species and was relatively unimportant in the factor analysis. Any effect on model performance either alone or by interaction with other site factors would likely be mediated through the substrate influencing soil temperature and hence the temperature experienced by weevils developing in the stumps. In the forecast modelling for this project, there were 16 peat-based sites and 11 mineral-based sites. The observed development was faster than expected suggesting that the temperature experience by the large pine weevil was higher than anticipated. Steep slope and mineral soil could result in very efficient drainage. So, soil would be dry and hence warm (despite being north facing, which should make it cooler) at a site such as Deerpark aforementioned. This can link to how site factors influence development. It is important to look at these anomalies to investigate the model efficacy. The depth of soils is important too. For larvae developing closer to the surface, especially in light soils, or in logs or fallen trees on the ground (Scott & King, 1974), temperature variations would be much large, resulting in greater acceleration in development rate (Inward et al., 2012). Soil type and slope may influence drainage and soil temperature. The effects of moisture conditions on temperature have been shown to influence soil microbiology and the aggregate stability of the soil. One study (Dowdesdale-Downey et al., 2023) found that for sandy loam soils for instance, aggregate stability decreased significantly with increasing moisture content. For a clay soil, aggregate stability increased significantly with increasing temperature. For both soil textures, temperature and moisture content affected microbiological community composition and respiration. This can show how varying soil types respond to temperature too.

It is worth noting too how pine weevils’ microclimates are affected or not when developing in soils. The developing weevils in the stump have minimal opportunity to change their location but there is potential movement when adult weevils may choose the location to lay eggs, or feeding weevils may slowly change their location within the stump as they tunnel through the wood (Fedderwitz *et al.* 2022). Hylobius larvae can develop in roots less than 10mm in diameter and may move through the soils if their food source is depleted. However, most larvae are found close to the body of the stump just below the soil level (Heritage & Moore, 2001).

Studies have shown that taxa responding to elevation (given the influence in this model) were indeed most strongly defined by temperature, as opposed to other environmental conditions with micro-topography affected taxa mainly controlled by temperature and soil moisture. This can show the microbial community composition is dependent on long term developments of near ground temperatures and soil moisture regimes with a certain resilience (Frindte *et al.* 2019). In terms of depth in the soil, the forecast model was shown to be very dependent on its "depth" argument input from the original model (Wainhouse *et al.* 2014). Sites like Hortland and Kilduff had a much better predicted curve when depth was equal to 1 and it got worse as this value increased. Interestingly, sites such as Lackenrea 1 and 2 had the opposite effect. Other sites such as Summerhill (co-located site) was not affected by changing the value of the variable. Other factors not investigated in the forecast model including time of felling. The time of felling can influence the time at which the parent generation of weevils colonise stumps and lay eggs and hence the initiation of development. (Moore et al., 2004; Teagasc, 2020).

5.3 Potential of "`pineR`" within the field and future work

The previous modelling in Maynooth by Williams et al (unpublished) aimed to inform management and improve management options for large pine weevil. This for example included indicating that fallow periods may be a viable future strategy to manage pine weevil which is not currently policy in Ireland but is practice in the UK. This research also laid the foundation for timing of plant protection products (PPPs) to be based on predicted emergence and voltinism patterns, as the current emergence model refines in terms of emergence. The `pineR` model proposes correcting the UK voltinism model to forecast first year emergence of large pine weevil at Irish clear fell sites. There is potential to expand to second year emergence in future work, but this was not covered in this research. The aim of this research that was to more fully validate a forecasting model using a more extensive dataset of Irish sites (27) presented with correction based on the 50% quartile in random forest. The stump hacking data in terms of weevil populations in-stump prior to emergence is only used for comparative purposes in my results but has not been integrated within `pineR`. It is proposed in the short term that `pineR` currently available on GitHub will be published on CRAN as a R programming package (Lemos dos Santos *et al*. 2023). Another potential avenue for future work is to use the population structure data to make it function within the forecast model.

One way to incorporate this would be to look the observed and simulated data for a particular site. If we had data for a given day of emergence (e.g., around day 200 for Glendine 2), we could integrate that as day 0 in a model or initiate a stage from this point in the model. Here the larvae and pupae are of most importance in terms of future emergence of adult weevils. With this data, a time cumulative model could be proposed that would take the proportion of pupae and use that as a predictor of how long it would take from now to first emergence, generating a simulated prediction. This facility would operate in a two-stage modelling framework where hacking data would be used to look at the proportion of pupae to find a prediction from when emergence is going to start. The model would then be used visually with for instance a curve matched to the predicted start date on another curve of the two stages of development. Adults could be also added to one value proportion of pupae if needed. A multinomial model covering larvae, pupae and adults could operate in this function too. Therefore, the potential model would operate with a proportion of larvae, pupae, and adults at one time point that can be used as a starting point to estimate or base a model on these proportions. This would bring us from the initial day to start of emergence and join the curves to create a more accurate forecast based on population structure information. The applicability of the `pineR` package and the machine learning corrected model in other geographical regions could also be investigated as well as different species. Additionally, the integration of other environmental variables may be further explored for model development (Lemos Dos Santos *et al.* 2023).

This work could also potentially lead to an R Shiny app developed that is easy to use and accessible for foresters on-site. The app could also be linked to potential policy applications in terms of monitoring weevil emergence. Further work could be explored based on UK modelling such as the Hylobius Management Support System (Willoughby *et al.*, 2020; Forest Research, 2021) and more recent approaches to assess weevil numbers on site using pheromones or other attractants and/ or the machine-vision-based Hylopod system. Another policy application from this research in terms of future work could also be low-cost weather stations on clear fell sites above 300 metres given the noted influence of altitude on the forecast model. While this research showed errors exist and due diligence should be done on any further development of the `pineR` forecast model, it displays the importance of site-specific covariates within the correction model and how such data can aid emergence forecast particularly first year emergence of large pine weevil. Further research could also adopt numerous approaches in the context of warming climate.

Williams et al (unpublished) showed that by the 2050s and 2080s climate projections, populations in Ireland will have an increased proportion of two-year pine weevil life cycles especially in the south where 100% of the population are projected to have two-year life cycles. Changes in voltinism reveal different temporal patterns emphasising the importance of modelling local, as well as regional scale variation in the effects of climate change in Europe. Further research would clearly require more in-depth evaluation on climate data in Ireland in the context of future emergence forecasts. In terms of future IPM for pine weevil, methods for various IPM categories are constantly developed both in Ireland and the UK such as population management tools, genotyping of tree species to be more tolerant, testing of alternative chemicals or more work on natural products or enemies such as nematodes (Willoughby *et al.,* 2020; Teagasc, 2020, McNamara *et al.,* 2018; Fedderwitz et al., 2022); accurate forecasting of weevil populations can have an important place in IPM of this pest. The implication of errors (i.e. multiple days error) is important though in establishing forecasting of the emergence and growth of the pest as incorrect timing could lead to issues with pest management if incorrect.

Ultimately, more knowledge and research are essential to provide more options and opportunities for managing this important forest pest. Indeed, also it is important to acknowledge that there will be likely changes in forest management in the future. This may affect efforts to model pine weevil, and the likelihood of damage. For example, as well as the rapid effects of climate change, sustainable forest management itself is likely to have more continuous cover forestry, and less clear-felling in the coming decades marking a rapid change if forestry policy (DAFM, 2023). In addition, mixed species stands will be much more common with native provenance too. This could have impacts on weevil development as this shift more towards diverse forests takes hold and less monoculture or Spruce stands become present with more and more multifunctioning forests. The changing climate equally will also impact the distribution of invasive species and pests. Therefore, moving forward, our forests must be resilient to cope with or even take advantage of growing conditions that a changing climate may bring. Forest Culture in Policy change is important too. For a variety of historical reasons, a culture where forestry is seen as an integral part of traditional agriculture has been slow to develop in Ireland. In certain parts of the country there is a longer tradition of forestry and greater acceptance and integration of it observing the benefits it can bring to local communities (Lyons, Coillte, *Pers. Comm.,* 2021). However, for many farmers, forestry is considered as an alternative land-use for their most marginal land and is not part of their identity in the same way, dairy, beef, or crop production may be in rural Ireland (DAFM, 2023).

5.4 Key Recommendations

5.4.1: Recommendations for Policy

This research hopes to lead to future work in the IPM developing forecasting methods for pest management. The new national forestry strategy offers interesting insights in terms of forest management in Ireland (DAFM, 2023). It is important to consider how this work can aid IPM practices in Ireland. We can observe the refining of the forecast model and how that can be used to help forestry policy and future developments. The DAFM forestry division require forest managers to prepare and submit forest management plans when a conifer forest reaches 10 years of age and is 10 Ha or greater in area. The plan must be prepared by a registered forester following a field assessment. This document includes tree stocking rates, nutrient status, average heights, yields and crucially when any management interventions are needed (Teagasc, 2021). It is important to understand the dynamics of pest populations for this reason and a forecasting tool in the right context can make significant difference. Nearly 11% of the land of Ireland is afforested and the most recent national policy has promoted afforestation to reach an adjust target of 18% by 2050 (DAFM, 2023). In Ireland, most forest stands are managed under clearfell, and replant systems and it is a condition of felling sites that sites must be replanted (Fedderwitz et al., 2022). The changing climate will also impact the health and productivity of our forests. This may include the frequency and intensity of things like forest fires and extreme weather events, but also the phenology in terms of seasonal timing of pest attacks. This has the potential to affect the adaptability of trees to climate change and may have social, economic, and environmental consequences in terms of reduced carbon sequestration. Our forests need resilience and even need to be able to take advantage of things like site specific factors in a future where growing conditions may change as the result of climate change. Forest expansion will require sustainable forest management and as the latest national forest strategy emphasises growing the right trees in the right places to promote carbon sequestration and storage as well as general forest health (DAFM, 2023). This strategy also includes measures such as measuring embodied carbon and promotion of substitution products.

Serious threats from forest pests and diseases are on the increase due to more globalisation, trade, and the impacts of climate change. This is compounded by climate-related abiotic affects such as fire, wind, and flooding. Ireland does not have the range of forest pests and disease that are endemic on the continent but has had significant outbreaks in recent years notably with ash dieback (*Hymenocyphus fraxineus)* in ash trees. Achieving favourable forest health status incorporates the need to ensure our forests are adaptable and resilient to the future climate.

Our forestry stakeholders are important in this regard and active involvement with constructive input is needed as a key enabler to develop fit for purpose polices and legislation in forest management. Ongoing monitoring and assessment of the health and condition of Irish forests can contribute to their overall resilience and biodiversity to aid wood production and ecosystem services (DAFM, 2023). Indeed, as many county councils set about their climate action plans, tree strategies are as important as ever in terms of things like carbon offsets in towns through shading for cool urban spaces but also in terms of the protection of trees. The Laois Heritage Forum recently mentioned this regarding the draft county climate action plan citing their joined-up strategy within the midlands region to review tree planting methods noting that the protection of existing tree stocks was as vital as future planting. It was noted how 2500 saplings carbon offset would be the equivalent of a 100-year-old tree to explain this point (Dempsey & Moore, *Pers. Comm.,* 2023). Including biodiversity indicators into forest planning is also growing in importance as it is a support for other ecosystem services. The overriding objective of the national forestry strategy was to expand our national forest estate radically and urgently on both public and private land in a manner that would deliver lasting benefits for climate change mitigation, biodiversity, sustainable wood production, economic development, employment, and the quality of life. The afforestation scheme will also aim to aid the rural economy and increase areas of native forests as well as improving purpose designed recreational areas. Equally benefits for water quality and the improvement and enhancement of biodiversity is important in the policy (DAFM, 2023). Research and innovation are key to realising projects like this and ambitious goals of forest strategies. There are barriers in terms of capacity and long-term research capacity. Forest research is currently performed by research performing organisations (RPOs) and funded through the departments competitive research funding programme. This has increased the quality of research in terms of more peer-reviewed publications and the variety of RPOs in the area. However, the very long nature of forest research means due to political, funding and even employment timeframes, this can last decades in terms of recording. This highlights the need for continuity and more flexible funding models with good archiving of both specific subject and data, and above all a focus on key factors derived from research that can guide future good silvicultural practice in land use, yields and impacts. This can lead to better controlling the establishment, growth, composition, health and quality of our future forests and woodlands (DAFM, 2023; Fedderwitz et al., 2022).

5.4.2: Recommendations for Further Research

The influence of site-specific factors such as elevation was shown on the model and perhaps low-cost weather stations on site as mentioned could be utilised as one policy recommendation to forecast weevil populations more accurately based on the temperature, they are experiencing on higher altitude sites for such future development of a forecast model incorporating temperature. Inexpensive temperature sensors could be further adopted as used in previous Irish and UK Research via Tinytag sensors on site or existing lab methods (Williams et al., unpublished, Inward et al., 2012). Satellite data could also be beneficial. The Spotta system offers 24/7 monitoring of large pine weevil as explained to the researcher on site in 2021 (Lyons, Coillte, *Pers. Comm.*, 2021). Improving data quality and availability is of course important to ensure the robustness of a forecast model. All factors could aid future weevil research. Utilising the current capabilities of weevil IPM in Ireland (Fedderwitz et al, 2022; Teagasc, 2020) and abroad is important to consider too in terms of things like The Hylobius management Support System (MSS) for instance (Forest Research, 2021). Its system ultimately has similar goals to this proposed forecast model to minimise transplant damage for foresters and aid operational goals while reducing chemical intervention and having a more sustainable forest management. terms of slope as a factor, influence was shown as well. One field study (Tolbert, 1975) looking at the effect of eight different aspects (slope exposures) in eastern Tennessee on arthropod distribution patterns (pitfall traps used for collection) found looking at various characteristics (available moisture, temperature, rainfall measurements, prevailing winds, solar radiation & vegetation cover) that southern and western exposures were categorised as warm and dry, while northern and eastern slopes were cool and wet. The southern facing slopes in that study were the most preferred for arthropod activity. The southeast and northwest slopes and the least arthropod activity and both were windswept. Glendine, a south facing aspect site was the most abundant site of weevil trial sites utilised within `pineR` in terms of emergence. Studies on Forest composition and community structures however have shown greater productivity on Northern aspects but dependant on location and temperature too (Harris, *et al.*, 2019; Solbreck & Gyldberg, 1979).

The `pineR` model had more sites than existing research in Ireland and aided the identification of mismatches between predicted and observed sites. The microhabitat distribution relative to the soil surface of developing weevils could be affected by aspect too. (McCutchan and Fox, 1986). Conceivably slopes could affect the model runs in terms of temperatures offsets as shown. Sites with offsets regarding model performance were important too. The recognition of differences in the gridded data correction for instance shows that the mechanism of transferring the model could have a possible shortcoming in terms of the 1km gridded model when sourcing a mean elevation over a high terrain (Walsh, 2012). Gortnapisha, as noted above, shows limitations in temperature constraints or the model been “too cold” to forecast emergence in its development. The comparison between random forest utilised in `pineR` to other potential methods could be further investigated in future work. It is proposed that further research should investigate different machine learning methods to improve performance of the corrected model, such as Bayesian additive regression trees, regularised regression, neural networks, and support vector machines. Machine learning offers a range of methods in terms of statistical theory to investigate to further improve the forecast modelling in `pineR` (Mello & Ponti, 2018). Issues such as the influence of depth were noted in research for improvement too in terms of better representations of forecasting first emergence. Soil was shown to have less influence on weevil activity as per above.

The economic value of the forest and maximisation of timber volumes has had more traditional focus in forest planning as opposed to the ecological value and presence of diverse habitats. (Teagasc, 2021). To forecast biodiversity potential, foresters can equally use models that simulate growth and other biological and ecological process. As models are simplified versions of reality, they may exclude important components of biodiversity’s multi scales and multi-facets. One study (Hunault-Fontbonne & Eyvindson, 2023) found that many biodiversity models used in forest planning mainly focus on structure and species elements, with less complexity and minor focus on connectivity and functions and more on genetic diversity. The inclusion of connectivity and biodiversity in future models is a possible recommendation to aid research further aiding phenological and biodiversity models. This research aims to lead to further development towards improving our knowledge of species adaptability and aiding improved IPM. The resourcing of forest genetic research may also aid this process to improve planting stock (Zas et al., 2017) and ensure suitably adapted reproductive material for our range of commercial and native forests to help the need for forestry establishment in Ireland. Forecast modelling offers potential always to aid Integrated Pest management.

If the role of the various factors which cause changes in abundance/emergence can be understood and related to predictable events, then modelling in its development can be properly managed to enable predictions for pest control for instance (Van Emden, 1989; Duffy et al., 2017). The potential of Information communication Technology for the semination of farmers/foresters has needs to be always realised further. IPM practices utilising such may be needed to solve many problems of sensitivity and intractability in the sustainable development of agriculture (Zuo Rui *et al.* 2012). In terms of climatology, the development of sophisticated regional climate models with high resolution can help improve pest simulation modelling in areas such as topography influenced factors and extremes (Flannagan *et al.* 2019). The results of the above forecast model in `pineR` demonstrate the potential of machine learning to improve upon traditional modelling techniques as evidenced by the favourable outcomes mostly when apply the corrected to model to Irish trial site weevil data (Lemos Dos Santos *et al.* 2023) and equally enhancing applicability in an Irish context from the existing UK model (Wainhouse *et al.*, 2014). `PineR` has a place in creating a forecasting template to be further developed with the forecast model to create a refined method for forecasting large pine weevil emergence in Ireland utilising domain specific knowledge.

5.5 Conclusion

The influence of temperature on biochemical reactions and resultant arthropod activity may either enhance or limit the effectiveness of integrated pest management. (Hallman & Denlinger, 1998). In the interest of optimising management including chemical intervention according to the principles of IPM if required at the core of this research, non-pesticide options have been discussed. These include early planting, mounding, feeding barriers and use of more vigorous and thicker girth planting stock to adapt to the challenges faced by weevil attack on site (Willoughby *et al.*, 2020; Teagasc, 2020; Fedderwitz et al., 2022). Cultural methods of plant protection have been investigated. These are practices that avoid high levels of pest infestation by developing conditions that are undesirable for the pest to survive or thrive. Site specific parameters can also help explain mismatches on site in terms of the forecast modelling between the observed and modelled data as seen in this research. Ultimately the covariates of interest can be used as identified to show factors that may significantly influence the model and take them into account to provide more reliable and accurate forecast predictions. This was seen notably with elevation being highly influential in terms of variable importance to certain sites and taken account of in the corrected forecast model. A crucial part of successful IPM against the large pine weevil has been able to estimate when the damage risk is highest to act or respond to emergence. Therefore, understanding how clearfell site factors influence pine weevil, with a view to emergence prediction in the case of the forecast model is essential. It is also important to improve understanding of pine weevil dynamics in terms of peak damage periods and how our changing climate and warmer temperatures may affect weevil development (Inward *et al.*, 2012) leading to potentially extended feeding periods. The `pineR` corrected forecast model forms an invaluable step in weevil emergence forecasting in Ireland. It creates a forecast model proposing a correction to an existing UK model (Wainhouse *et al.*, 2014) that utilises clearfell geographic data matched to temperature data that drives specifically more accurate and reliable predictions of large pine weevil first year emergence at sites in Ireland. The step lays a path for a future forecasting tool that will refine the forecasting method of this important forest pest in Ireland.

Bibliography:

Abroi.D and Shankar.U (2012) *Integrated Pest Management: Principals and Practice*: *History, Overview and principals of Ecologically based pest management*, Published by C.A.B International. pp. 1-27.

Aizen MA, Patterson WA (1995) Leaf phenology and herbivory along a temperature gradient - a spatial test of the phenological window hypothesis. Journal of Vegetation Science, 4, 543-550.

Alexandridis, N. Marion, G. Chaplin-Kramer, R. Dainese, M., Ekroos, J., Grab, H., Jonsson, M., Karp, D.S., Meyer, C., O'Rourke, M.E, Pontarp, M., Poveda, K., Seppelt, R., Smith, H.G, Martin, E.A, Clough, Y. (2021), Models of natural pest control: Towards predictions across agricultural landscapes, *Biological Control*, Volume 163, 104761.

Allen, I.C. (1976), A modified sine wave method for calculating degree days. *Environmental Entomology*, 5, pp. 388-396.

Altermatt, F. (2010) Climatic warming increases voltinism in European butterflies and moths. Proceedings. Biological sciences / The Royal Society, 277(1685), 1281– 1287.

Angilletta, M., Dunham, A. (2003). The Temperature-Size Rule in Ectotherms: Simple Evolutionary Explanations May Not Be General. *The American naturalist*. 162. 332-42.

Arthurs, S., Thomas, M.B., 2001. Effect of dose, pre-mortem host incubation temperature and thermal behaviour on host mortality, mycosis, and sporulation of Metarhizium anisopliae var. acridum in Schistocerca gregaria. *Biocontrol Science and Technology*, 11, 411–420.

Ayres, M.P. & Lombardero, M.J. (2000) Assessing the consequences of global change for forest disturbance from herbivores and pathogens. *Science of the Total Environment*, 262, pp. 263–286.

Bachmetjew P (1900) Freezing of body fluids in insects. *Zeitschrift für wissenschaftlichen Zoologie* 67: 529–550

Bailey J.E., Johnson G.L., Toth S.J. (1994) Evolution of weather-based peanut leaf spot spray advisory in North Carolina. *Plant Dis.*, 78, pp. 530-535

Bale, J.S. (1999) Impacts of Climate warming on artic aphids: A comparative analysis. *Ecological bulletins* 47; pp. 38-47.

Bale, J. S. *et al.*. (2002) 'Herbivory in global climate change research: Direct effects of rising temperature on insect herbivores', *Global Change Biology*, 8(1), pp. 1–16. doi: 10.1046/j.1365-2486.2002.00451.x

Bebber, Daniel & Ramotowski, Mark & Gurr, Sarah. (2013). Crop pests and pathogens move poleward in a warming world. *Nature* Climate Change. 3.

Bejer-Petersen, B. (1975). Length of development and survival of Hylobius abietis as influenced by silvicultural exposure to sunlight. *Arsskrift, Kongelige Veterinaer- og Landbohoejskole, Koebenhavn.*, pp. 111-120.

Bentz BJ, Logan JA & Amman GD (1991) Temperature-depen- dent development of the mountain pine beetle (Coleoptera: Scolytidae) and simulation of its phenology. Canadian Entomologist 123: 1083–1094

Birch, L.C. (1957) The role of weather in determining the distribution and abundance of animals. *Cold Spring Harbor Symposia on Quantitative Biology,* 22, pp. 203–218.

Boote, K., Jones, J. & Pickering, N. (1996). Potential Uses and Limitations of Crop Models. *Agronomy Journal*. 88. 704-716.

Bradshaw CJ, Leroy B, Bellard C, Roiz D, Albert C, Fournier A, Barbet-Massin M, Salles JM, Simard F, Courchamp F. (2016) Massive yet grossly underestimated global costs of invasive insects. *Nat Commun*. Oct 4; 7:12986.

Brown, J.H, Gillooly. J.F, Allen, A.P, Savage.V.M, West, G.B (2004) **Toward a metabolic theory of ecology** Ecology, 85 (2004), pp. 1771-1789

Buckley, Y., Finn, J., Matthews, A., Moran, J. & Stout, J. (2020). *The Common Agriculture Policy post-2020: A new Green Architecture,* Technical Report, Novel Eco-schemes, and biodiversity indicators.

Campbell, A., Frazer, B., Gilbert, N., Gutierrez, A. & Mackauer, M. (1974) Temperature requirements of some aphids and their parasites. Journal of Applied Ecology, 11, 431–438.

Carson, R. (1962) *Silent spring*, Houghton Mifflin, Boston, Massachusetts.

Carton, Y., Réaumur (1683–1757): the discoverer of parasitoids in France, *Biological Control,* Volume 32, Issue 1, 2005, Pages 40-48.

Ceballos, G., Ehrlich, P.R., Dirzo, R., 2017. Biological annihilation via the ongoing sixth mass extinction signalled by vertebrate population losses and declines. Proc. Natl. *Acad. Sci.* 114, E6089–E6090

Cesaraccio, C., Spano, D. , Duce, P. & Snyder, R. (2001). An improved model for determining degree-day values from daily temperature data. International journal of biometeorology. 45. 161-9.

Chuine. I, de Cortazar-Atauri, I.G, Kramer.K, Hänninen.H (2013) **Plant development Models_2013** M.D. Schwartz (Ed.), Phenology: An Integrative Environmental Science, Springer, Netherlands, Dordrecht (2013), pp. 275-293.

COFORD (2003) National Council for Forest Research and Development, Ted Horgan, Michael Keane, Richard McCarthy, Michael Lally, and David Thompson, Edited by Joe O' Carroll. First Published by Coford in 2003. Belfield, Dublin 4.

Colinet, H., Sinclair, B., Vernon, P. & Renault, D. (2014). Insects in Fluctuating Thermal Environments. Annual review of entomology. 60.

Coulson, R.N. and Saunders, M.C. (1987) Computer assisted decision making as applied to entomology. *Annual review of entomology,* 32, pp.415-437.

DAFM (2015) *Foodwise 2025*, Growth Opportunities, Forestry section, pp 38. Published by Department of Agriculture, fisheries and the marine and the government of Irleand.

DAFM (2023) *Ireland's Forestry Strategy 2023-2030,* Section 3 (Stategic drivers for change), Section 5 (Objectives, Values & Goals) and Section 6(Strategic enablers), pp. 17, 32 &43. Published by Department of Agriculture, Fisheries and the marine and the governemnt of Ireland. Published 6th September 2023.

Danks, H.V. (1987) Insect dormancy: an ecological perspective 1st ed. H. V. Danks, ed., Biological Survey of Canada (Terrestrial Artropods) (Ottawa

Dawidziuk A., Kaczmarek J., Jedryczka M. (2012) The effect of winter weather conditions on the ability of pseudothecia of *Leptosphaeria maculans* and *L. biglobosa* to release ascospores. *Eur. J. Plant Pathol.*, 134, (2), pp. 329-343

Della Sala, D. A. and Goldstein, M. I. (2017) *The anthropocene climate: Humanity on a planetary collision course*, *Encyclopedia of the Anthropocene*. Elsevier Inc. doi: 10.1016/B978-0-12-809665-9.10536-1.

Dempsey, S. & Moore, K. (2023), *Pers.Comm,* Laois Heritage Forum, held in Portlaoise Library on September 26th 2023. Presentation of county climate action plan draft.

Denlinger, D.L and Lee Jnr., R.E. (1998) Ch.3 Physiology of cold sensitivity, *In* Hallman, G.J. and Denlinger, D.L. (Edited by) (1998) *Temperature Sensitivity in Insects and Application in Integrated Pest Management,* 2nd Edition, Routeledge, New York. Pp. 55-97.

Dent, D. (1991), *Insect Pest Management,* Published by C.A.B International. Oxford.

Dent, D. (1995), *Integrated Pest Management*, Published by Springer Science & Business Media, Oxford.

Dent, D. (2000), *Insect pest management,* 2nd edition, Published by CABI bioscience, UK centre, Ascot, UK.

De Roode JC, Lefèvre T. (2012) Behavioral Immunity in Insects. *Insects*.3(3):789-820. Published 2012 Aug 15.

Dillon, A.B., Ward, D., Downes, M.J., Griffin, C.T., 2006. Suppression of the large pine weevil Hylobius abietis (L.) (Coleoptera: Curculionidae) in pine stumps by entomopathogenic nematodes with different foraging strategies. Biol. Control 38, 217–226.

Dillon, A., Griffin, C., 2008. Controlling the large pine weevil, Hylobius abietis, using natural enemies. Silviculture/Management No. 15. Coford Connects, Dublin.

Dillon, A.B., Moore, C.P., Downes, M.J., Griffin, C.T., (2008) Evict or infect? Managing populations of the large pine weevil, *Hylobius abietis*, using a bottom–up and top–down approach, *Forest Ecology and Management,* Volume 255, Issue 7, pp. 2634-2642,

Dirzo, R. *et al.* (2014) 'Defaunation in the Anthropocene', *Science*, 345(6195), pp. 401–406. doi: 10.1126/science.1251817.

Drake, V.A., Wang, H.K. and Harman, I.T. (2002) Insect monitoring radar: remote and network operation. *Computers and electronics in Agriculture*, 35, 77-94.

Duffy, C., Fealy, R. and Fealy, R. M. (2017) ‘An improved simulation model to describe the temperature-dependent population dynamics of the grain aphid, Sitobion avenae’, *Ecological Modelling*. Elsevier B.V., 354, pp. 140–171.

Dunlap. T (2008). *DDT, Silent Spring and the rise of environmentalism*, Edited by Thomas R. Dunlap (2008) University of Washington press.

Donatelli, M., Magarey, R.D., Bregaglio, S., Willocquet, L., Whish, J.P.M., Savary, S. (2017) Modelling the impacts of pests and diseases on agricultural systems, *Agricultural Systems*, Volume 155, pp. 213-224.

Dowdeswell-Downey, E., Grabowski, R.C. & Rickson, R.J. (2023) Do temperature and moisture conditions impact soil microbiology and aggregate stability? *J Soils Sediments* **23**, 3706–3719 (2023). https://doi.org/10.1007/s11368-023-03628-2

ECAD.EU (2021), E-OBS gridded climate dataset (V.24.1), Mean temperature data for Europe obtained from open access European observations by European Climate assessment dataset.

Elliot, S.L., Blanford, S., Thomas, M.B., 2002. Host–pathogen interactions in a varying environment: temperature, behavioural fever and fitness. Proc. R. Soc. Lond. B 269, 1599–1607.

Elliot, S.L., Horton, C.M., Blanford, M., Thomas, M.B., 2005. Impacts of fever on locust life-history traits: costs or benefits? Biol. Lett. 1, 181–184. http://dx.doi.org/ 10.1098/rsbl.2004.0279

Esker P.D., Savary S., McRoberts N. (2012) *CAB Reviews: perspectives in Agriculture, Veterinary Science, Nutrition and Natural Resources.* 052. CAB Reviews 2012: Crop loss analysis and global food supply: focusing now on required harvests; 7, pp. 1–14.

Fajvan MA and Wood JM. (1996) Stand structure and development after gypsy moth defoliation in the Appalachian Plateau. *Forest Ecol Manag* **89**: 79–88.

FAO (2009) Global review of forest pest and diseases: A thematic study prepared in the framework of the global forest resources assessment 2005: *FAO forestry paper*, 156. Food and agriculture organisation of the United Nations, Rome, 2009.

FAO (2012), *OECD-FAO Agricultural Outlook 2012*, OECD Publishing, Paris pp.75

Fedderwitz, F. Griffin, C.T. and McNamara, L. (2022). *Integrated pest management of the large pine weevil Hylobius abietis: review and perspectives for Ireland.* Forestry Research Section, Teagasc.

Filazzola.A, Matter,S.F., Scott MacIvor. J (2021), The direct and indirect effects of extreme climate events on insects,*Science of The Total Environment,* Volume 769,

Fishel.F.M. (2016) *Pest management and pesticides: A historical perspective*, UF/IFAS Extension: University of Florida, PI 219

Fitt, G.P., Dillon, M.L. and Hamilton, J.G. (1995) Spatial dynamics of Helicoverpa populations in Australia: simulation modeling and empirical studies of adult movement. Computers and Electronics in Agriculture 13, 177–192

Fitt.G, and Wilson.L.(2012) *Integrated Pest Management: Principals and Practice: IPM for Sustainable Agriculture*, CABI International. pp. 27-41.

Flanagan, J., Nolan, P., McGrath, R. & Werner, C. (2019). Towards a definitive historical high-resolution climate dataset for Ireland – promoting climate research in Ireland. Advances in Science and Research. 15. 263-276. 10.5194/asr-15-263-2019.

Frenne, P., Lenoir, J., Luoto, M., Scheffers, B., Zellweger, F., Aalto, J., Ashcroft, M., Christiansen, D., Decocq, G., De Pauw, K., Govaert, S., Greiser, C., Gril, E., Hampe, A., Jucker, T., Klinges, D., Koelemeijer, Irena., Lembrechts, J., Marrec, R. & Hylander, K. (2021). Forest microclimates and climate change: Importance, drivers, and future research agenda. Global Change Biology. 27.

Friedenberg, N.A., Sarkar, S., Kouchoukos, N., Billings, R.F., Ayres, M.P., 2008. Temperature extremes, density dependence, and southern pine beetle (Coleop- tera: Curculionidae) population dynamics in east Texas. Environmental Ento- mology 37 (3), 650–659.

Frindte K, Pape R, Werner K, Löffler J, Knief C. (2019) Temperature and soil moisture control microbial community composition in an arctic-alpine ecosystem along elevational and micro-topographic gradients. *ISME* J. 2019 Aug;13(8):2031-2043.

Forest Research. (2021). Hylobius Management Support System (MSS). Retrieved 15th September 2021 from https://www.forestresearch.gov.uk/tools-andresources/fthr/tree-health-and-protection-services/hylobius-management-supportsystem/.

Forrest, J.RK. (2016) Complex responses of insect phenology to climate change, *Current Opinion in Insect Science*, Volume 17, 49-54.

Forzieri, Giovanni & Girardello, Marco & Cecch erini, Guido & Spinoni, Jonathan & Feyen, Luc & Hartmann, Henrik & Beck, Pieter & Camps-Valls, Gustau & Chirici, Gherado & Mauri, Achille & Cescatti, Alessandro. (2021). Emergent vulnerability to climate-driven disturbances in European forests. Nature Communications. 12. 1081.

Foster, S.P., Devine, G., Devonshire, A.L., 2007. Insecticide resistance. In: Van Emden, H.F., Harrington, R. (Eds.), Aphids as Crop Pests. CAB International, Oxfordshire, pp. 261–286.

Georghiou, G.P. (1990) Overview of insecticide resistance: In: Green, M.B., LeBaron, H.M. and Moberg, W.K. (eds), *Managing resistance to agrochemicals*. American Chemical society, Washington, DC, pp. 18-41.

Golizadeh A & Zalucki MP (2012) Estimating temperature-dependent developmental rates of potato tuberworm, *Phthorimaea operculella* (Lepidoptera: Gelechiidae). *Insect Science* 19: 609– 620.

Gonzalez-Tokman, D., Córdoba-Aguilar, A., Dáttilo, W., Lira-Noriega, A., Sánchez Guillén, R. & Villalobos, F. (2020). Insect responses to heat: physiological mechanisms, evolution, and ecological implications in a warming world. *Biological Reviews*. 95.

Gordon, I., Calatayud, P.-A., Le Gall, P. & Garnery, L. (2019) *We are losing the "Little things that run the world".* United Nations Environment Foresight Brief 011.

Gordo, O. & Sanz, J.J. (2005) Phenology and climate change: a long-term study in a Mediterranean locality. Oecologia, 146(3), 484–495.

Hall, R. (1995) Challenges and prospects of integrated pest management, In.Reuveni, R. (Ed.) , *Novel approaches to Integrated pest management,* Lewis publishers, Boca Raton, Florida, pp. 1-9.

Hallman, G.J. and Denlinger, D.L. (Edited by) (1998) *Temperature Sensitivity in Insects and Application in Integrated Pest Management,* 2nd Edition, Routeledge, New York.

Harp, P.R. (2005), Cypermethrin: Environmental Fate, Editor(s): Philip Wexler, *Encyclopedia of Toxicology* (Second Edition), Elsevier, pp. 714-716.

Harris, J. E., Rodenhouse, N. L. and Holmes, R. T. (2019) 'Decline in beetle abundance and diversity in an intact temperate forest linked to climate warming', *Biological Conservation*. Elsevier, 240(August), p. 108219. doi: 10.1016/j.biocon.2019.108219.

Haydock, L.A.J. , Pomroy, W., Stevenson, M.A., Lawrence, Ke. (2016). A growing degree-day model for determination of Fasciola hepatica infection risk in New Zealand with future predictions using climate change models. Veterinary Parasitology. 228.

Heinrich, B. (1993) *The hot blooded Insects,* Harvard University Press, Cambridge.

Heinrich, B. (1973). Mechanisms of Insect Thermoregulation. In: Wieser, W. (eds) Effects of Temperature on Ectothermic Organisms. Springer, Berlin, Heidelberg

Heritage, S. & Moore, R. (2001) The assessment of site characteristics as part of a management strategy to reduce damage by *Hylobius, Forrestry commission information note,* Published April 2001.

Hickling, R., Roy, D.B., Hill, J.K., Fox, R. & Thomas, C.D. (2006) The distributions of a wide range of taxonomic groups are expanding polewards. *Global Change Biology*, 12(3), 450–455.

Himmer. A. (1932) Die Temperaturverhiiltnisse bei den sozialen Hymenopteren *Bioi. Rev.* 7(3): 224-253

Hirao, T.M., Murakami, M. and Kashizaki, A. (2008) "Effects of mobility on daily attraction to light traps: comparison between lepidopteran and coleopteran communities. *Insect conservation and Diversity.* 1, pp.32-39.

Hodkinson ID, Bird J (1998) Host speci®c insect herbivores as sensors of climate change in Arctic and alpine environments. Arctic and Alpine Research, 30, 78-83

Hodkinson, Ian. (2005). Terrestrial insects along elevation gradients: Species and community responses to altitude. Biological reviews of the Cambridge Philosophical Society. 80. 489-513.

Honek, A. 1999. Constraints on thermal requirements for insect development. Entomol. Sci. 2: 615-621.

Hultin, E. (1955) Influence of temperature on the rate of enzymatic processes. *Acta Chemica Scandinavica* 9, 1700.

Hunault-Fontbonne, J & Eyvindson, K. (2023) Bridging the gap between forest planning and ecology in biodiversity forecasts: A review, Ecological Indicators, Volume 154, 2023, 110620, ISSN 1470-160X, https://doi.org/10.1016/j.ecolind.2023.110620.

Hunter MD (1992) A variable insect plant interaction ± the relationship between tree budburst phenology and popula- tion levels of insect herbivores among trees. Ecological Entomology, 17, 91-95.

Inward, D. J. G., Wainhouse, D. and Peace, A. (2012) 'The effect of temperature on the development and life cycle regulation of the pine weevil Hylobius abietis and the potential impacts of climate change', *Agricultural and Forest Entomology,* 14(4), pp. 348–357.

IPCC (2013) Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change (IPCC). Stocker, T.F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S.K., Boschung, J., Nauels, A., Xia, Y., Bex, V. & Midgley, P.M. (Eds.). Cambridge University Press, Cambridge, UK and New York, NY, USA, 1535 pp.

Janes JK, Li Y, Keeling CI, *et al.* (2014) How the mountain pine beetle (*Dendroctonus ponderosae*) breached the Canadian Rocky Mountains. *Mol Biol Evol* **31**: 1803–15.

Johnson, F.H. & Lewin, I. (1946) The growth rate of E. Coli in relation to temperature, quinine, and coenzyme. Journal of cellular and comparative physiology, 28, pp. 47-75.

Johnston, I.A., Bennett, A.F., 1996. *Animals and Temperature: phenotypic and Evolutionary Adaptation*. Cambridge University Press, Cambridge.

Jonsson, A.M., Appelberg, G., Harding, S., Barring, L., 2009. Spatio-temporal impact of climate change on the activity and voltinism of the spruce bark beetle, Ips typographus. Global Change Biology 15 (2), 486–499.

Jones, P. D. and Osborn, T. J. (2018) *Climate Record: Surface Temperature Trends , Reference Module in Earth Systems and Environmental Sciences*. Elsevier Inc. doi: 10.1016/b978-0-12-409548-9.11292-8.

Kapranas, A., Malone, B., Quinn, S., Mc Namara, L., Williams, C.D., O'Tuama, P., Peters, A., Griffin C.T. (2017) Efficacy of entomopathogenic nematodes for control of large pine weevil, *Hylobius abietis*: effects of soil type, pest density and spatial distribution. *Journal of pest science*, 90: pp. 495–505.

Keating, B.A., Carberry, P.S., Bindraban, P.S., Asseng, S., Meinke, H. and Dixon, J. (2010). Eco-efficient agriculture: Concepts, challenges and opportunities. *Crop Science*, 50, S109-S119.

Kleuken, A.M., Hau, B., Ulber, B. and Poehling, H.M. (2009) Forecasting migration of cereal aphids (Hemiptera: Aphididae) in autumn and spring. *Journal of Applied Entomology*, 133, pp.328-344.

Knutson, A.E. and Muegge, M.A. (2010) A degree-day model initiated by pheromone trap captures for managing pecan nut casebearer (Lepidoptera: Pyralidae) in pecans. *Journal of Economic Entomology*, 103, pp. 735-743.

Kogen, M. and Turnipseed, S.G. (1987) Ecology and management of soybean arthropods. *Annual Review of Entomology*, 32, pp. 507-538.

Koul, O., Dhaliwal, G.S. and Cuperus, G.W. (2004) Edited by, *Integrated Pest Management: Potential Constraints and Challenges*. C.A.B.I Publishing. Oxofrd U.K.

Krogh, A., & Zeuthen, E. (1941). The Mechanism of Flight Preparation in Some Insects. *The Journal of Experimental Biology, 18*, 1-10.

Lactin, D. J., N. J. Holliday, D. L. Johnson, and R. Craigen. 1995. Improved rate model of temperature-dependent development by arthropods. Environ. Entomol. 24: 68-75.

Lalík, M., Galko, J., Kunca, A., Nikolov, C., Rell, S., Zúbrik, M., Dubec, M., Vakula, J. ,Gubka, A., Leontovyč, R., Longauerova, V., Konôpka, B., Holusa, J. (2021). Ecology, management and damage by the large pine weevil (Hylobius abietis) (Coleoptera: Curculionidae) in coniferous forests within Europe. *Central European Forestry Journal*. 67. 91-107. 10.2478/forj-2021-0005.

Langstrom, B., Day, K.R., (2004) Damage control and management of weevil pests, especially *Hylobius abietis*. In: Lieutier, F., Day, K.R., Battisti, A. Gregoire, J.C., Evans, H.F. (Eds.), *Bark and Wood Boring Insects in Living Trees in Europe, a Synthesis*. Kluwer Academic Publishers, Dordrecht, Netherlands, pp. 415–444

Leather, S., Day, K. & Salisbury, A. (1999). The biology and ecology of the large pine weevil, *Hylobius Abietis* (Coleoptera: Curculionidae): A problem of dispersal? *Bulletin of entomological research*, 89(1), pp. 3-16.

Lehmann, P., Ammunet, T., Barton, M., Battisti, A., Eigenbrode, S.D., Uhd. Jepsen, J., Kalinkat, G., Neuvonen, S., Niemela, P., Terblanche, J.S., Okland, B., Bjorkman, C. (2020) Complex responses of global insects' pests to climate warming. *Frontiers in Ecology and the Environment*, 18(3), pp. 141-150.

Lemos dos Santos, A., *et al.* (2023), *Machine Learning Corrected life cycle model for the large pine weevil*, based on pineR, Hamilton Institute, Department of Biology & Department of Geography, Maynooth University. Package available on CRAN-R. https://github.com/Alessandra23/pineR.

Liebhold, A.M. and Tobin, P.C. (2008) Population ecology of insect invasions and their management. *Annual review of Entomology*, 53, pp. 387-408.

Logan, J.A. and Weber, L.A. (1989) Population Model Design System (PMDS): User's Guide. Department of Entomology, Virginia Polytechnic Institute and State University, Blackburg, Virginia

Lyons.C (2021), *Pers.Comm*, Online meeting on 5th October 2021 with Colm Lyons, Coillte Silvacultural specialist and advisor on project.

Ma, Z.S. and Bechinski, E.J. (2008) A survival-analysis-based simulation model for Russian wheat aphid population dynamics. Ecological Modelling 216, 323–332.

Mattsson, Anders. (2016). Reforestation challenges in Scandinavia. *REFORESTA*. 1. 67.

McCain, C.M, Garfinkel, C.F. (2021), Climate change and elevational range shifts in insects, *Current Opinion in Insect Science*, Volume 47, Pages 111-118,

Maxwell, S., Fuller, R., Brooks, T. *et al.* Biodiversity: The ravages of guns, nets, and bulldozers. *Nature* **536,** 143–145 (2016)

Medina-Pastor, Paula & Triacchini, Giuseppe. (2020). The 2018 European Union report on pesticide residues in food. EFSA Journal. 18. 10.2903/j.efsa.2020.6057.

Mello, R. & Ponti, M. (2018). Machine Learning: A Practical Approach on the Statistical Learning Theory. 10.1007/978-3-319-94989-5.

Menzel, A., Sparks, T., Estrella, N., Koch, E., Aasa, A., Ahas, R., Alm-Kubler, K., Bissolli, P., Braslavska, O., Briede, A., Chmielewski, F., Črepinšek, Z., Curnel, Y., Defila, C., Donnelly, A., Filella, I., Jatczak, K., Mage, F., Zust, A. (2006). European phenological response to climate change matches the warming pattern. *Global Change Biology*. 12. pp.1969-1976

Merril, S.C., Gebre-Amlak, A., Armstrong, J.S. and Pearirs, F.B. (2010) Nonlinear degree-day models of the sunflower weevil (Curculionidae: Coleoptera). Journal of Economic Entomology, 103, pp. 303-307.

Metcalf, R.L. (1980) Changing role of insecticides in crop protection. *Annual review of Entomology*, 25, pp. 215-250.

Metcalf, R.L (1996), Applied Entomology in the Twenty-First Century, *American Entomologist*, Volume 42, Issue 4, Winter 1996, pp. 216–227.

Met Eireann (2021), Met Eireann online website, Climate section, Available data, *Historical data from current stations, Hourly, daily, and Monthly data resolution.* Available from: https://www.met.ie/climate/available-data/historical-data Accessed on: 06th October 2021

Met.ie (2022), Met Eireann Website, Climate Section, *"What we measure – Met Eireann's Weather Observation Network"*. Available from: https://www.met.ie/climate/what-we-measure [Accessed on: 8th February 2022]

Moore, R. (2001). Emergence trap developed to capture adult large pine weevil *Hylobius abietis* (Coleoptera: Curculionidae) and its parasite *Bracon hylobii* (Hymenoptera: Braconidae). *Bulletin of Entomological Research*, *91*, 109-115.

Moore, R., Brixey, J. & Milner, AD. (2004). Effect of time of year on the development of immature stages of the Large Pine Weevil (*Hylobius abietis* L.) in stumps of Sitka spruce (*Picea sitchensis* Carr.) and influence of felling date on their growth, density, and distribution. *Journal of Applied Entomology*. 128. 167 - 176. 10.1111/j.1439-0418.2004.00828. x.

Mora. C, Rollo. A, Tittensor, D.P., Science 443, 295 (2013).

Morris, R.F. and Miller, C.A. (1954) The development of life tables for the spruce budworm. Canadian Journal of Zoology, 32, pp. 283-301.

Müller, C. B., & Schmid-Hempel, P. (1993). Exploitation of cold temperature as defence against parasitoids in bumblebees. *Nature, 363*(6424), 65–67

Munro, J.W. (1928) The biology a control of *Hylobius abietis* L. Forestry 2, 31–39

Munro, H.L. , Montes, C.R., Kamal, J.K., Gandhi, M. and Poisson, A. (2022) A comparison of presence-only analytical techniques and their application in forest pest modeling, *Ecological Informatics*, Volume 68, 101525,

Murray, T. (2019) *The silent extinction of Insects*. Online article. Biodiversity Ireland. Available from: https://biodiversityireland.ie/the-silent-extinction-of-insects/. Accessed online [20th March 2021).

Netherer, S. and Schopf, A. (2010) 'Potential effects of climate change on insect herbivores in European forests-General aspects and the pine processionary moth as specific example', *Forest Ecology and Management*, 259(4), pp. 831–838. doi: 10.1016/j.foreco.2009.07.034

Nilsson, U., & Örlander, G. (1995). Effects of regeneration methods on drought damage to newly planted Norway spruce seedlings. *Canadian Journal of Forest Research-Revue Canadienne De Recherche Forestiere, 25*, pp. 790-802

Newman J.A., Gibson D.J., Parsons A.J., Thornley J.H.M. (2003) How predictable are aphid population responses to elevated CO2. *Journal of animal ecology*, **52**: pp. 556–566.

Nokes S.E., Young J.H. (1991) Simulation of the temporal spread of leafspot and the effect on peanut growth *Trans. ASAE,* 34, pp. 653-662

Nordenhem, H. & Nordlander, G. (1994) Olfactory oriented migration through soil by root living *Hylobius abietis* (L.) larvae (Col., Curculionidae). Journal of Applied Entomology 117, 457–462.

Nordlander, G., Eidmann, H.E., Jacobsson, U., Nordenhem, H., Sjodin, K., (1986) Orientation of the pine weevil *Hylobius abietis* to underground sources of host volatiles. *Entomol. Exp. Appl.* 41, 91–100.

Olatinwo R.O., Paz J.O., Brown S.L., Kemerait R.C., Culbreath A.K., Beasley J.P. Jr, *et al.* (2008). A predictive model for spotted wilt epidemics in peanut based on local weather conditions and the *Tomato spotted wilt virus* risk index.
Phytopathology, 98, pp. 1066-1074.

Olatinwo, R., Hoogenboom, G. (2014) *Chapter 4 - Weather-based Pest Forecasting for Efficient Crop Protection*, Editor(s): Dharam P. Abrol, Integrated Pest Management, Academic Press, pp. 59-78.

O'Tuama.P (2021) (Former Silvicultural Officer with Coillte and Maynooth Research trials), Personal communication on phone interview on the 2nd of December 2021.

Oerke, E.C., Dehne, H.W., Schönbeck, F. and Weber, A. (1994) *Crop production and crop protection – Estimated crop losses in major food and cash crops*, Elsevier, The Netherlands.

Oerke, EC. (2006) Crop losses to pests. *Journal of agricultural science.* 144 pp.31-43

Oerke, E C (2007). "Crop Losses to Animal Pests, Plant Pathogens, and Weeds", In: David Pimental (ed.), Encyclopaedia of Pest Management, 116-120, CRC Press, USA.

Oliveira, C.M., Auad, A.M, Mendes, S.M., Frizzas, M.R. (2014)
Crop losses and the economic impact of insect pests on Brazilian agriculture,
Crop Protection, Volume 56, 2014, pp. 50-54,

Örlander, G., & Nilsson, U. (1999). Effect of reforestation methods on pine weevil (*Hylobius abietis*) damage and seedling survival. *Scandinavian Journal of Forest Research, 14*, pp. 341-354.

Örlander, G., Nordlander, G., & Wallertz, K. (2001). Extra food supply decreases damage by the pine weevil *Hylobius abietis*. *Scandinavian Journal of Forest Research, 16*, pp. 450-454.

Oreskes, Naomi & Shrader-Frechette, Kristin & Belitz, Kenneth. (1994). Verification, Validation, and Confirmation of Numerical Models in the Earth Science. Science (New York, N.Y.). 263. pp. 641- 646.

Parmesan, C. & Yohe, G. (2003) A globally coherent fingerprint of climate change impacts across natural systems. Nature, 421(6918), 37–42.

Pasquali.S, Soresina.C, Gilioli.G (2019) The effects of fecundity, mortality and distribution of the initial condition in phenological models, *Ecological Modelling,* Volume 402, 2019, Pages 45-58.

Pathak, M.D. (1968) Ecology of common insect pests of rice. *Annual review of Entomology*, 13, pp. 257-294.

Payne NM (1929) Absolute humidity as a factor in insect cold hardiness with a note on the effect of nutrition on cold hardiness. *Annals of the Entomological Society of America* 22: 601–620.

Petersson, M., Örlander, G., & Nordlander, G. (2005). Soil features affecting damage to conifer seedlings by the pine weevil *Hylobius abietis*. *Forestry: An International Journal of Forest Research, 78,* 83-92.

Pinkard, E.A., Battaglia, M., Bruce, J., Leriche, A. & Kriticos, D.J. (2010) Process-based modelling of the severity and impact of foliar pest attack on eucalypt plantation productivity under current and future climates. Forest Ecology and Management, 259, 839–847

Pinkard.E, Battaglia.M, Bruce.J, Matthews.S, Callister.A.N, Hetherington.S, Last.I, Mathieson.S, Mitchell.C, Mohammed.C, Musk.R, Ravenwood.I, Rombouts.J, Stone.C, Wardlaw.T (2014), A history of forestry management responses to climatic variability and their current relevance for developing climate change adaptation strategies, *Forestry: An International Journal of Forest Research*, Volume 88, Issue 2, April 2015, Pages 155–171.

Prasad, Y. G. and Prabhakar, M. (2012) 'Pest monitoring and forecasting', *Integrated Pest Management: Principles and practice*, (May), pp. 41–57

Pollard, C.P., Griffin, C.T. Moral, R., Duffy, C., Chuche, J., Gaffney, M.T., Fealy, R.M, Fealy, R. (2020) phenModel: A temperature-dependent phenology/voltinism model for an herbivorous insect incorporating facultative diapause and budburst, Ecological Modelling, Volume 416, 2020, 108910.

Poveda, K., Martínez, E., Kersch-Becker, M.F., Bonilla, M.A. & Tscharntke.T (2012)
Landscape simplification and altitude affect biodiversity, herbivory, and Andean potato yield J. Appl. Ecol., 49., pp. 513-522.

Pureswaran, D. S., Roques, A. and Battisti, A. (2018) 'Forest insects and climate change', *Current Forestry Reports*. Current Forestry Reports, 4(2), pp. 35–50. doi: 10.1007/s40725-018-0075-6.

Raffa, K.F., Aukema, B.H., Bentz B.J, Carroll, A.L., Hicke, J.A., Turner, M.G. and Romme, W.H. (2008) Cross-scale drivers of natural disturbances prone to anthropogenic amplification: the dynamics of bark beetle eruptions, *Bioscience*, 58, pp. 501-517

Rebaudo, F. and Rabhi, V. B. (2018) 'Modeling temperature-dependent development rate and phenology in insects: review of major developments, challenges, and future directions', *Entomologia Experimentalis et Applicata*, 166(8), pp. 607–617. doi: 10.1111/eea.12693.

Refaeilzadeh, P., Tang, L. & Liu, H. (2009). Cross-Validation. Encyclopedia of Database Systems. 532–538. 532-538. 10.1007/978-0-387-39940-9_565.

Regniere, J., 2009. Predicting insect continental distributions from species physiol- ogy. , Unasylva No. 231/232 (vol. 60/1–2), pp. 37–42.

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

(RMetS,2020), Rmets.org, climate section, article, *Climate modelling: How is earth's climate modelled and how modelling help our understanding*. Published on 15th December 2020. Available online: https://www.rmets.org/metmatters/climate-modelling Accessed: 20th January 2021.

Ritchie, J.C. (1989) Current thoughts on Community Organisation, *Ecological society of America*, 70, (2) April 1989, pp. 525-526.

Rose, D., Leather, S.R., Matthews, G.A., (2005) Recognition and avoidance of insecticide-treated Scot's pine (*Pinus sylvestris*) by *Hylobius abietis* (Coleoptera: Curculionidae): implications for pest management strategies. *Agr. Forest Entomol*. 7, 187–191.

Rosenzweig, C., Iglesias, A., Yang, X.B., Epstein, P.R., Chivian E. (2001)
Climate change and extreme weather events. Implications for food production, plant diseases, and pests. *Glob. Chang. Hum. Health*, 2, pp. 90-104

Roy, D.B. & Sparks, T.H. (2000) Phenology of British butterflies and climate change. Global Change Biology, 6(4), 407–416.

Sánchez-Bayo, F. and Wyckhuys, K. A. G. (2019) 'Worldwide decline of the entomofauna: A review of its drivers', *Biological Conservation*, 232(January), pp. 8–27. doi: 10.1016/j.biocon.2019.01.020.

Savary, S., Teng, P.S., Willocquet, L., Nutter Jr., F.W. (2006) Quantification and modelling of crop losses: a review of purposes. *Annu. Rev. Phytopathol.*, 44, pp. 89-112

Scheffers. B.R, Joppa. L.N., Pimm.S.L., Laurance. W.F, Trends Ecol. Evol. 27, 501–510 (2012).

Schoolfield, R.M., Sharpe, P.J. and Magnuson, C.E. (1981) Non-Linear regression of Biologically dependent rate models based on absolute reaction-rate theory. *Journal of theoretical biology*, 88, pp. 719-731.

Schoonhoven LM, van Loon JJA, Dicke M (2005) Insect-Plant Biology, Ed 2. Oxford University Press, Oxford

Scott, T.M. & King, C.J. (1974) The large pine weevil and black pine beetles. *Forestry Commission Leaflet*, 58, 1–23

Sehnal, F. (1991) Effects of cold on morphogenesis, *In* R.E.LEE and D.L. Denlinger eds., *Insects at low temperature.* Pp. 149-173, Chapman and Hall, New York.

Sharpe, J.H. and DeMichele, D.W. (1977) Reaction kinetics of poikilotherm development. *Journal of theoretical Biology*, 64, pp. 649-670.

Skourti, A., Kavallieratos, N. G. and Papanikolaou, N. E. (2019) 'Laboratory evaluation of development and survival of Tribolium castaneum (Herbst) (Coleoptera: Tenebrionidae) under constant temperatures', *Journal of Stored Products Research*. Elsevier Ltd, 83, pp. 305–310. doi: 10.1016/j.jspr.2019.07.009.

Solbreck, C. (1980). Dispersal distances of migrating pine weevils, *Hylobius abietis*, Coleoptera: Curculionidae. *Entomologia Experimentalis Et Applicata, 28*, pp. 123-131.

Solbreck, C., & Gyldberg, B. (1979). Temporal flight pattern of the large pine weevil, *Hylobius abietis* L. (Coleoptera, Curculionidae), with special reference to the influence of weather. *Zeitschrift fur Angewandte Entomologie, 88,* pp. 532-536.

Southwood, T.R.E. and Way, M.J. (1970) Ecological background to pest management, in concepts of pest management, (eds R.L. Rabb and F.E. Guthrie), North Carolina State University Press, Raleigh, pp. 6-29.

Sparks, T.H. and Woiwod, I.P. (1999) Dates of insect appearance and activity. In: Cannell, M.G.R., Palutikof, J.P. and Sparks, T.H. (eds) Indicators of Climate Change in the UK. Centre for Ecology and Hydrology, UK, 58–59.

Sporleder, M., Simon, R., Gonzales, J., Chavez, D., Juarez, H., De Mendiburu, F. and Krosche, J. (2009). ILCYM – Insect life cycle modelling: A software package for developing temperature-based insect phenology models with applications for regional and global risk assessments and mapping. International Potato Centre, Lima, Peru.

Stefanescu, C., Penuelas, J. & Filella, I. (2003) Effects of climatic change on the phenology of butterflies in the northwest Mediterranean Basin. Global Change Biology, 9(10), 1494–1506.

Strand, J.F. (2000) Some agrometeorological aspects of pest and disease management for the 21st Century. *Agricultural and Forest Meteorology,* 103, pp. 73-82.

Strathdee, A., Bale, J., Block, W., Coulson, S., Hodkinson, I. & Webb, N. (1993). Effects of temperature elevation on a field population of Acyrthosiphon svalbardicum (Hemiptera: Aphididae) on Spitsbergen. *Oecologia*. 96. 457-465.

Strathdee AT, Bale JS (1995) Factors limiting the distribution of Acyrthosiphon svalbardicum (Hemiptera: Aphididae) on Spitsbergen. Polar Biology, 15, 375-380

Stoakley, J.T. & Heritage, S.G. (1990) Application leaflet on the use of 'Permit' and 'Permasect 25 EC' for pre-planting treatment of young trees against *Hylobius abietis* and Hylastes spp. *Research Information Note 177,* Farnham, Forestry Commission Research Division

Sunday, J., Bates, A., Kearney, M., Colwell, R., Dulvy, N. , Longino, J. & Huey, R. (2014). Thermal-safety margins and the necessity of thermoregulatory behavior across latitude and elevation. Proceedings of the National Academy of Sciences of the United States of America. 111.

Tan, J. Y., Wainhouse, D., Morgan, G., & Day, K. R. (2011). Interaction between flight, reproductive development, and oviposition in the pine weevil Hylobius abietis. *Agricultural and Forest Entomology, 13,* pp. 149-156

Tarr, S.A.J. (1972) Principles of plant pathology, Macmillan, London

Taylor, L.R. (1963) Analysis of the effect of temperature on insects in flight. *Journal of animal ecology* 32, 99-117.

Teagasc (2017) Teagasc Forestry Development Department , 'Reforestation following Clearfell Reforestation', *Farm Foresry Series* , April 2017 (20).

Teagasc (2020) Teagasc Forestry development departmenet, "Manageing Pine Weevil: Methods and Options, *Farm forestry Series,* September 2020 (24)

Teagasc (2021), Harvesting and Selling Timber from Conifer Forests: A Step-by-Step Guide for Forest Owners, *Harvesting and Selling Timber series*, Published by Teagasc Forestry Development Section, Mellows Campus, Athenry, Co. Galway, H65R718, Published December 2021.

Thomas, M & Blanford, S. (2003). Thermal biology in insect-parasite interactions. Trends in Ecology & Evolution. 18. 344-350.

Thomas, C.D. (2010) Climate, climate change and range boundaries. Diversity and Distributions, 16(3), 488–495.

Tischner, H. (2000) Instruments of the warning service for plant protection in Bayern (Germany). EPPO Bulletin 30, 103–104.

Tivy, J. (1973) *Biogeography: A study of plants in the Ecosphere.* Second Edition. Oliver and Boyd. Croythorn House, 23 Ravelston Terrace, Edinburgh. pp. 133-157.

Tolbert, W. W. (1975). The Effects of Slope Exposure on Arthropod Distribution Patterns. *The American Midland Naturalist*, *94*(1), 38–53. https://doi.org/10.2307/2424536

Tonnang, H.EZ., Sokame, B.M, Abdel-Rahman, E.M, Dubois, T. (2022) Measuring and modelling crop yield losses due to invasive insect pests under climate change, *Current Opinion in Insect Science,* Volume 50,100873.

Trudgill, DL ; Honek, A; Li, D; Van Straalen, NM (2005), Thermal time - concepts and utility *ANNALS OF APPLIED BIOLOGY, pp.* 146,1-14 DOI10.1111/j.1744-7348.2005.04088.x

Uelmen, J. A. *et al.* (2016) 'Effects of winter temperatures, spring degree-day accumulation, and insect population source on phenological synchrony between forest tent caterpillar and host trees', *Forest Ecology and Management*. Elsevier B.V., 362, pp. 241–250.

Van Emden, H.F. (1989) *Pest Control.* 2nd Edition, Chapman and Hall, New York. pp. 27, pp 99.

Van Emden, H.F. and Harrington, R. (2007) *Aphids as Crop Pests* Published by C.A.B International, UK. .

Varley, G.C.and Gradwell, G.R. (1970) Recent advances in insect population dynamics. *Annual review of Entomology,* 15, pp.1-24.

Velarde, R.A.M., Wiedenmann, R.N. & Voegtlin, D.J. (2002) Influence of photoperiod on the overwintering induction of Galerucella calmariensis L. BioControl, 47(5), 587– 601.

Villamor. A.L, Carreño. S, López-Goldar. X, Suárez-Vidal. E, Sampedro. L, Nordlander. G, Björklund.N, Zas.R (2019) Risk of damage by the pine weevil Hylobius abietis in southern Europe: Effects of silvicultural and landscape factors, *Forest Ecology and Management*, Volume 444, Pages 290-298.

Visser. M.E, Both.C (2005) Shifts in phenology due to global climate change: the need for a yardstick. *Proceedings of the Royal Society B,* 272 (2005), pp. 2561-2569

Vitousek, P., Mooney, H., Lubchenco, J. & Melillo, J. (1997). Human Domination of Earth's Ecosystems. *Science*. 277.

Von Sydow, F. (1997). Abundance of pine weevils (*Hylobius abietis*) and damage to conifer seedlings in relation to silvicultural practices. *Scandinavian Journal of Forest Research, 12*, pp. 157-167

Wagner, T.L., Wu, H.I., Sharpe, P.J.H., Schoolfield, R.M. and Coulson, R.N. (1984), Modelling insect development rates: a literature review application of a biophysical model, *Annals of the Entomological society of America*, 77, pp. 208 – 225.

Wagner, T.L., Wu, H.I., Feldman, R.M., Sharpe, P.J.H. and Coulson, R.N. (1985) Multiple-cohort approach for simulating development of insect populations under variable temperatures. *Annals of the entomological society of America*, 78, 691-704.

Wainhouse, D., Brough, S., Greenacre, B. (2007) Managing the pine weevil on lowland pine, *Forest commision practice note*, Forest Research UK, June 2007.

Wainhouse, D., Inward, D. J. G. and Morgan, G. (2014) 'Modelling geographical variation in voltinism of Hylobius abietis under climate change and implications for management', *Agricultural and Forest Entomology*, 16(2), pp. 136–146. doi: 10.1111/afe.12043.

Wallertz, K., Örlander, G., & Luoranen, J. (2005). Damage by pine weevil *Hylobius abietis* to conifer seedlings after shelterwood removal. *Scandinavian Journal of Forest Research, 20*, pp. 412-420.

Wallner, W.E. (2000). *Lymantria dispar*. NAFC (North American Forest Commission) -ExFor Pest Report.

Wermelinger, B., & Seifert, M. (1998). Analysis of the temperature dependent development of the spruce bark beetle Ips typographus (L.) (Col., Scolytidae). *Journal of Applied Entomology, 122*(4), 185-191.

Williams, C.D., Dillon, A.B., Harvey, C.D., Hennessy, R., Mc Namara, L., Griffin, C.T. (2013) Control of a major pest of forestry, Hylobius abietis, with entomopathogenic nematodes and fungi using eradicant and prophylactic strategies. *Forest ecology management,* 305: pp. 212–222.

Willoughby, I., Moore, R., & Nisbet, T. (2017). Forest Research U.K., *Interim guidance on the integrated management of Hylobius abietis in UK forestry*.

Willoughby, I, , Moore, R, Moffat,A, Forster, J., Sayyed, I. & Leslie, K. (2020). Are there viable alternatives to the use of conventional insecticides for the protection of young trees from damage by the large pine weevil Hylobius abietis L. in UK forestry?. Forestry: An International Journal of Forest Research. 93. 10.1093/forestry/cpaa013.

Wojda, I. (2017) ‘Temperature stress and insect immunity’, *Journal of Thermal Biology*. Elsevier Ltd, 68(September 2016), pp. 96–103. doi: 10.1016/j.jtherbio.2016.12.002.

WOW.ie (2022), Weather Observations Website for Met Eireann, Live map of weather observations in Ireland as part of global WOW network of crowdsourced weather observations. Available from: https://wow.met.ie/ [Accessed on 12th February 2022]

Yang X.B., Scherm H. (1997) El Niño and infectious disease, *Science*, 275, p. 739.

Young, H. S. *et al.* (2016) ‘Patterns, Causes, and Consequences of Anthropocene Defaunation’, *Annual Review of Ecology, Evolution, and Systematics*, 47(August), pp. 333–358. doi: 10.1146/annurev-ecolsys-112414-054142.

Yin, X., Kropff, M.J., McLaren.G.. Visperas, R.M. (1995) A nonlinear model for crop development as a function of temperature, Agricultural and Forest Meteorology, Volume 77, Issues 1–2, PP. 1-16,

Ying, W., Xiaoning, L., Xiaohui, J. (2006) Difference between automatic and manual meteorological observations, *Environmental science*, 2006.

Yurk, B.P, Powell.J. A (2010) Modelling the effects of developmental variation on insect phenology. *Bulletin of Mathematical Biology*, 76 (2010), pp. 1334-1360

Zalom, F., Goodell, P., Wilson, L., Barnett, W.W. & Bentley, WJ. (1983). *Degree-Days, the Calculation and Use of Heat Units in Pest Management:* University of California Division of Agriculture and Natural Resources Leaflet 21373. Univ. of California, Berkeley.

Zalucki, M.P. and Furlong, M.J. (2005) Forecasting *Helicoverpa* populations in Australia: a comparison of regression-based models and a bioclimatic based modelling approach. *Insect Science*, 12, pp. 45-46.

Zas, R., Björklund, N., Sampedro, L., Hellqvist, C., Karlsson, B.et al., 2017: Genetic variation in resistance of Norway spruce seedlings to damage by the pine weevil Hylobius abietis. *Tree Genetics and Genomes*, 13:111.

Zhou, X.R, Harrington, I.P. Woiwod, *et al. (1995) Effects of temperature on aphid phenology.* Glob. Change Biol., 1 (1995), pp. 303-313

Zhou, Z.-S., Guo, J.-Y., Chen, H.-S. & Wan, F.H. (2010) Effect of humidity on the development and fecundity of Ophraella communa (Coleoptera: *Chrysomelidae*). BioControl, 55(2), 313–319.

Zumr, V., & Stary, P. (1994). Monitoring of seasonal occurrence of *Hylobius abietis* (L) (Col, Curculionidae) in different forst environments of a model area. *Journal of Applied Entomology-Zeitschrift Fur Angewandte Entomologie, 118*, pp. 361-364.

ZuoRui, S., DengFa, C., YunHui, Z. & LingWang, G. (2012). Role of information and communication technology in integrated pest management. *Integrated Pest Management: Principles and practice*, (May), pp. 433 – 439.

Appendices:

The following sections were completed with the assistance of Alessandra Lemos Dos Santos, main author of the `PineR` package and referenced as such throughout text (Lemos Dos Santos *et al.*, 2023). It looks at the intricate functioning of the `pineR` package. Additional graphics were not utilised in main text but are included in a final section from the whole thesis.

A(i) Functioning of `pineR` (Lemos Dos Santos *et al.* 2023)

The functioning of the `pineR` package ultimately has a principal function called indeed `"pineR"`. The arguments include the data input covering the variables of `"day"`, `"month"`, `"year"` `"min_temp"`, `"max_temp"` and `"temp"` that represents the day, month, year, minimum, maximum, and average temperatures respectively. With that data correctly adjusted and modified for use in R Code, the `npop`, `ntimes`, `ngen`, `species` and `depth` arguments are specified corresponding to population size, number of models runs, number of generations, tree species selected (Pine or spruce) and depth of stump or soil at clear-fell site respectively. The soil depth is inputted with the options 1, 2 or 3 corresponding to 10cm, 20cm and 30cm respectively. It is worth noting here there was observed soil and stump temperature data at Hortland and Summerhill that became available to researchers late in the assemblage of `pineR` which was not used in the coding of the model. The principal function it is worth noting is divided in the following internal functions which correspond to the previously mention life cycle stages of the simulation model. These include the `getStumpTemp`, `GetFirstGeneration` or internal functions within `module 1, 2,3,4, 5` and `6` internal functions of the overall forecasting package. These numbered modules refer to the modularised approach discussed extensively above of life cycle progression where there are internal functions implementing each of them. The principal output of the package is generated as `mature` (maturation feeding), `lavalst` (Larval stage), `gstmonth` (the oviposition month for the start of final generation), `geetime` (number of days to start of oviposition), `gemwinter` (number of winters to emergence above ground), `gemmonth` (month of emergence above ground), `stageov` (number of days from completion of overwintering to start of oviposition) and `rweevil` (Overwintering status of weevil). These outputs in code ultimately correspond the emergence simulation and predictive capabilities at trial sites generated from the `pineR` package.

A (II)– Random Forests

Random Forests (RF) is an ensemble machine learning algorithm that is used widely in classification and regression problems. This method builds decision trees by creating a multitude of decision trees at training time. These build on different sample data and take a majority vote for classification and average in case of regression. In the implementation of Random Forest, certain variables are described. In the first instance Let $Y_i \in \mathbb{R}\ (i = 1, \ldots, n)$ be an outcome variable. Then $\boldsymbol{x}_i$ is set as a p-dimensional vector of features, and D_n (a set containing observations) is expressed as $D_n = \{(\boldsymbol{x}_1, y_1), \ldots, (\boldsymbol{x}_n, y_n)\}$. This method was first introduced by Breiman (2001). RF is a non-parametric machine learning algorithm, which can be used for classification and regression problems, based on an ensemble of decision trees stated as $h(\boldsymbol{x}_i, ; \theta_m),\ m = 1, \ldots, M,$ where θ_m are independent and identically distributed (iid) random vectors, and M is the number of trees in the ensemble.

RF uses the Classification and Regression Tree (CART) algorithm (Breiman *et al.*, 1984) to build a set of (approximately) independent trees using bootstrap sampling. As in the bagging method (Bootstrap aggregating) (Breiman, 2001), RF uses different bootstrapped subsets which are randomly drawn with replacement from the data. However, in RF each decision tree uses a subset of features randomly selected from the original vector of features $\boldsymbol{x}_i$. To generate the final fit, the predictions from each decision tree are averaged, $\hat{h}(\boldsymbol{x}_i;\ \boldsymbol{\theta}) = \left(\frac{1}{M}\right)\sum_m h(\boldsymbol{x}_i, ; \theta_m)$. A basic example of the RF method for regression is illustrated in figure 4.11 below.

In terms of predictive power, RF is generally superior to parametric models, such as generalised linear/additive models, in terms of predictive performance and capabilities (Breiman, 2001) on a broad range of data sets as it does not impose or assume any parametric form for relation between the response variable and covariates. Additionally, RF is generally less computationally expensive than other competing machine learning algorithms (such as Bayesian Additive Regression Trees) and has the advantage of an in-built measure of variable importance (See section 4.3.6.3 for more details). However, a disadvantage of RF is the black-box nature of the predictions which can come with some interpretational issues (Goldstein *et al.*, 2015). Additionally, for RF regression, the algorithm is incapable of extrapolating values outside of the range of the response (Inglis *et al.*, 2022).

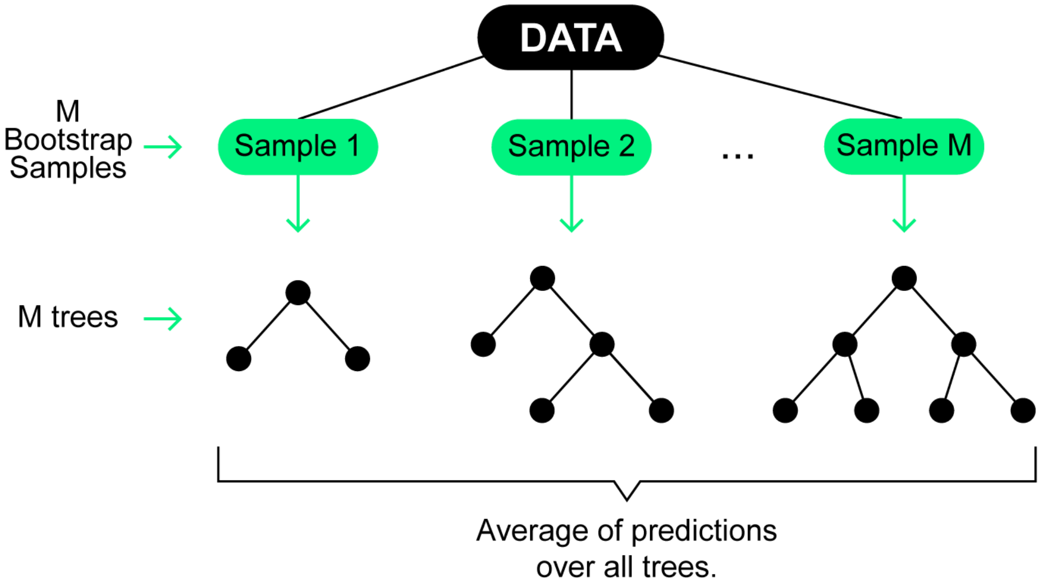


Figure 1: A flowchart representation describing the Random Forest algorithm for regression. The data is separated into *M* bootstrap samples and from each sample a tree is generated. For the final fit, the predictions from each tree are averaged.

A (III) Variable Importance

Variable importance (Vimp.) is a method used in machine learning to determine the relative ranking of importance of the variables used in constructing a model and gives an indication which variables are useful or relevant for predicting the response. Random Forests use a variety of techniques to determine the importance of a variable in forecasts. The most popular Vimp. method for regression is a permutation method which aims to calculate the reduction of some given metric, usually the mean square error (MSE).

Permutation importance was introduced by Breiman (2001) and is measured by observing the change in the model's predictive performance after a variable has been randomly permuted. Commonly, the percentage increase in MSE is used as a metric for estimating importance in regression models. This is achieved by initially noting the MSE for the model, then, randomly permuting each variable and comparing the MSE for the new dataset with the original. The resulting Vimp score is taken as the difference between the original model's MSE and the permuted model's MSE when a single variable value is randomly permuted. In Random Forests, the percentage increase in MSE is calculated by permuting the values of the out of bag (OAB) samples.

This permutation based MSE measure can be described as follows from (Grömping, 2009): The out of bag MSE (OOBMSE) is calculated, for a tree *t*, as the average of the squared deviations of OOB responses from their respective predictions:

$$OOBMSE_t = \frac{1}{n_{OOB,t}} \sum_{i \in OOB_t}^{n} (y_i - \hat{y}_{it})^2, \qquad (1)$$

In this formula, it is shown where $\hat{y}_{it}$ are the predictions for tree *t* associated with OOB observation i, and $n_{OOB,t}$ is the number of OOB observations in tree *t*. The idea from this is that if the predictor variable $x_j, j = 1, \ldots, p$, is unimportant in predicting *y*, then randomly permuting the values x_j, in the OOB data, this should have little to no effect on the prediction of *y*. Therefore, the value of the equation (2) for the $OOBMSE_t(\tilde{x}_j)$, where $\tilde{x}_j$ is the permuted value of variable x_j,this should not be considerably larger than $OOBMSE_t$.

$$OOBMSE_t(\tilde{x}_j) = \frac{1}{n_{OOB,t}} \sum_{i \in OOB_t}^{n} \left(y_i - \hat{y}_{it}(\tilde{x}_j)\right)^2. \qquad (2)$$

The difference $OOBMSE_t(\tilde{x}_j) - OOBMSE_t$ is then calculated, for each variable x_j in each tree *t* and is then averaged over all trees in the forest and normalised by the standard deviation of the differences. The result provides the percentage increase in the MSE for each predictor variable when compared to the MSE when that predictor is permuted.

A(IV) - Leave One Out Cross Validation

Cross-validation (CV) is a class of techniques to evolve the prediction performance of a model given a dataset by subsetting it into two groups: one to train the model and the other to validate the correction model in this case. In CV, different partitions can be made, depending on the data and the type of validation desired. Among the objectives when performing cross-validation, the ability of a model prediction or forecast can be evaluated, compared with the performance of one or more different models and determined the best model fit, as well as having control over the calibration of predictive distributions. A popular method for performing CV is the Leave-One-Out-Cross-Validation (LOOCV), a particular case of *k*-fold cross validation, where the number of folds and the number of instances events are equal. In LOOCV, it is not necessary to have a random partition, because every fold has just one instance (Wong, 2015).

This means that in each iteration of the algorithm, all the data is used to train the model, except for a single observation. Although LOOCV produces precision estimates with a high variance, these are unbiased. Also, LOOCV is a good choice when the number of instances in a dataset is small as in the case of this research with 27 emergence trial sites as said. For more details on this method, see (Wong, 2015), (Refaeilzadeh *et al*., 2009) and (Efron, 1983).

A(V) – Implementation of final correction forecast model.

In this section, all the methods previously mentioned in the last sections were used to correct the existing difference in the forecast predictions from the adapted Wainhouse *et al.* (2014) model and the observed emergence in each trial site in Ireland. In essence, the strategy employed was as follows:

1. Run the model (`pineR` implementation of Wainhouse *et al.* (2014)) with Irish sites.

2. Calculate the difference diff_{jq}, j = 1, …, 27, q = 0.25, 0.5 and 0.75 (days offset in model)

3. Create a dataset with 27 rows and 6 columns, where the first column corresponds to the response variable, and the other columns correspond to observations of the vector of covariates, for example $\boldsymbol{x_i} = [altitude, \text{slope}, \text{soil type}, \text{aspect}, \text{species}]$.

4. Obtain the predictions of the RF algorithm, considering the LOOCV (Leave one out cross validation).

5. Obtain the VImp. measures to identify covariates of interest.

6. Add to the value of the quantile q of the distribution obtained in (1) the predicted value of the difference calculated in (4). (Refer to above)

7. Plot the corrected CDF curve and its quantile interval to implement the finalised correction model.

The final corrections were based on the difference observed in the 0.5 (50%) quantile which was the best fit. However, for some sites the best corrections were observed for the 0.25 (25%) or 0.75 (75%) quantile values. Sites where this occurs, and unique results of the correction model will be discussed in the appropriate section (see the Results chapter). The quantile intervals were obtained considering 50 model replicates at emergence trial sites. The RF model was fit using the ranger algorithm within the Caret package (Kuhn, 2009) in R. The difference in emergence was used as the response and a LOOCV approach was used for the implementation. The model was tuned to optimise the RMSE value by selecting the MTRY value (That is, the depth of each tree was restricted), and the number of trees strictly defined was 100. To establish which variables are important for the forecast model determining the VImp. measures, the package VIVID (Inglis *et al.*, 2021) was used. The importance type was Permutation, and the size of the grid for evaluating the predictions was set to 100 to achieve this and verify the correction model.

A(Vi) – Additional results and graphics

```
dist1 <- 22.7 (Ballinla House)
dist2 <- 30.2 (Derrygreenagh)
dist3 <- 19.3 (Dunsany)

total_dist <- dist1 + dist2 + dist3.

w1_raw <- total_dist - dist1
w2_raw <- total_dist - dist2
w3_raw <- total_dist - dist3
total_weight <- w1_raw + w2_raw + w3_raw

w1 <- w1_raw / total_weight
w2 <- w2_raw / total_weight
w3 <- w3_raw / total_weight

weighted_temp <- site1$max_temp * w1 + site2$max_temp
* w2 + site3$max_temp * w3.
```

Figure A1: An example of the original calculation of the weighted temperature values for Hortland Trial site (2010-2012) Inputted in the model in R code.

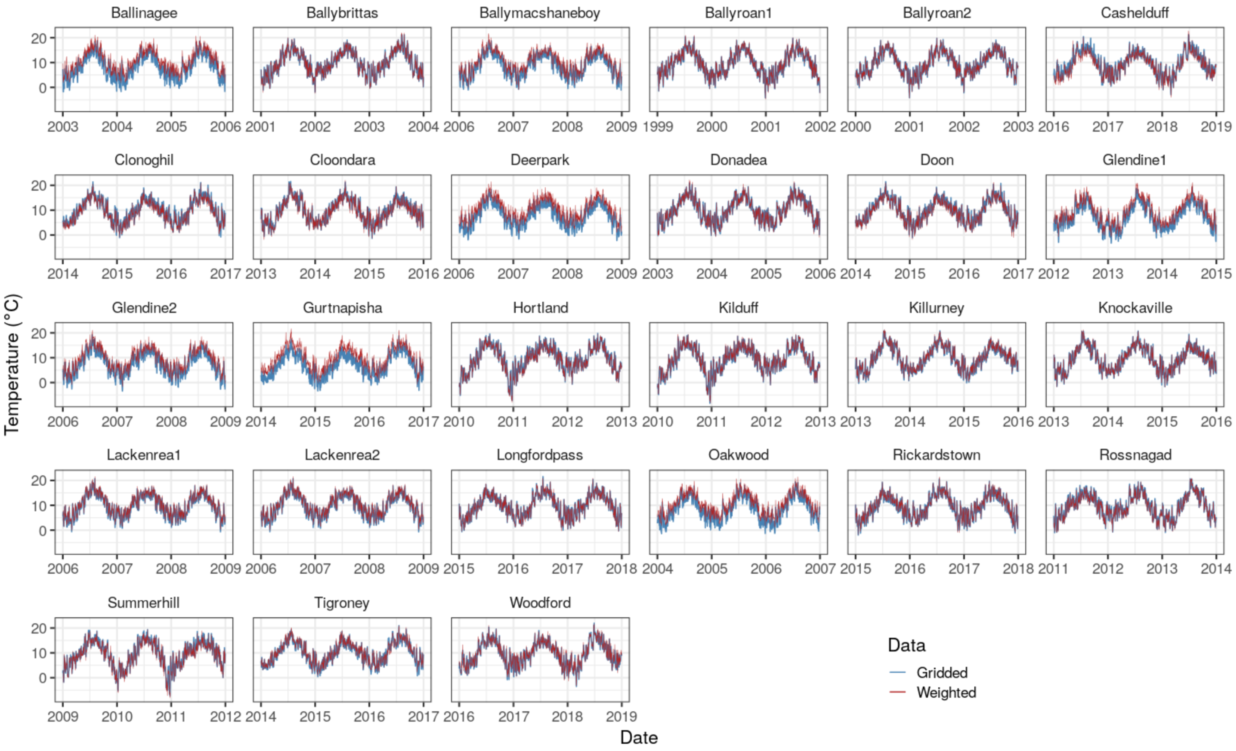
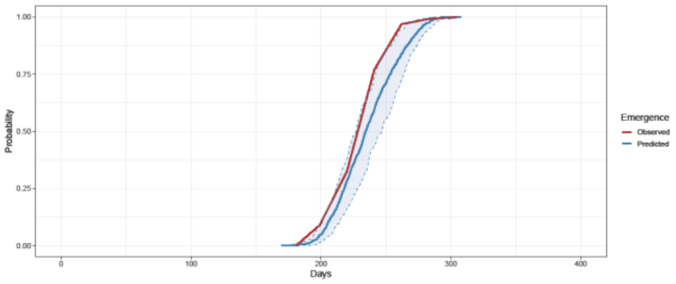
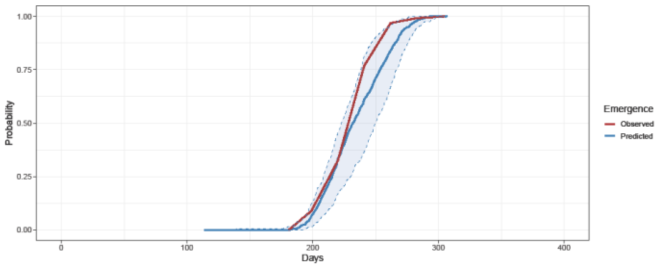
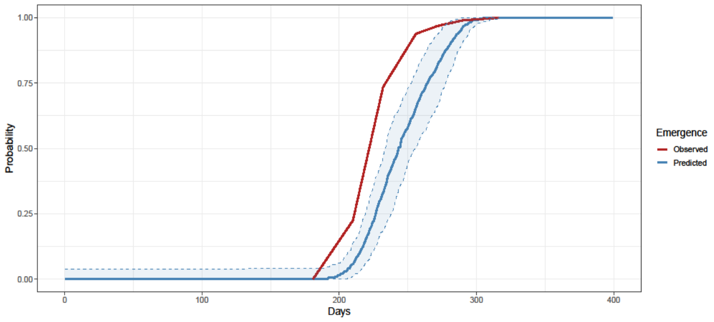
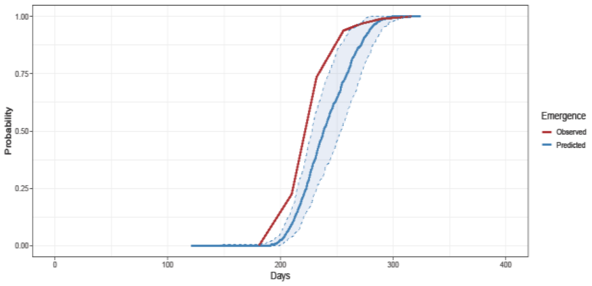
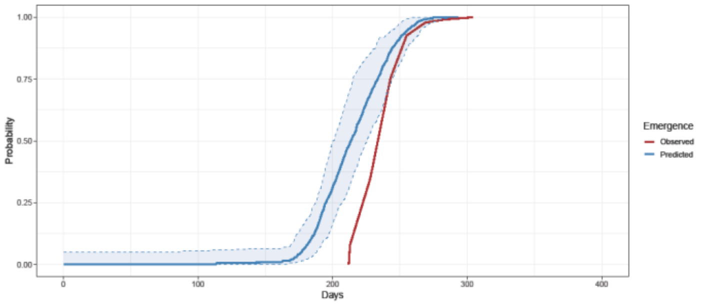
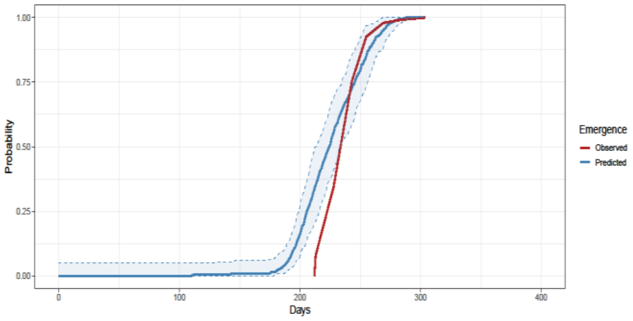
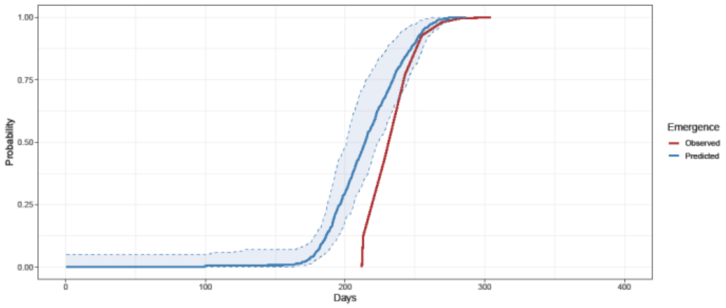
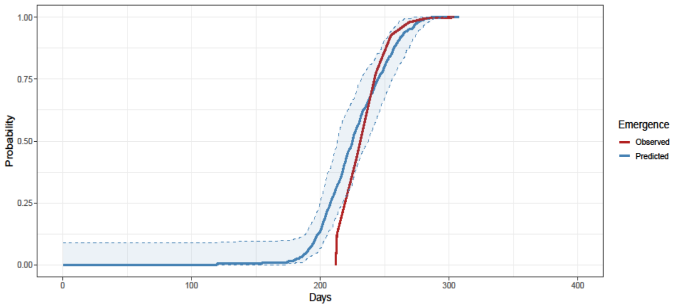


Figure A2: Comparison of gridded and weighted temperature data for each trial site utilised in the modelling of `pineR`.

Table A1: Comparison of observed weighted and gridded data temperature values for the 27 emergence trial sites utilised in the forecast modelling of `pineR` with use of the 0.01-degree resolution (1km) Met Eireann grids.

site	DG	DW
Ballinagee	-14	43
Ballybrittas	30	33
Ballymacshaneboy	19	58
Ballyroan1	-16	-15
Ballyroan2	9	12
Cashelduff	25	-22
Clonoghil	8	-1
Cloondara	15	12
Deerpark	-90	2
Donadea	-9	-5
Doon	-6	-14
Glendine1	-21	36
Glendine2	-46	28
Hortland	-14	-15
Kilduff	-31	-11
Killurney	19	27
Knockaville	-2	-3
ballyrao1	1	15
Lackenrea2	3	13
Longfordpass	39	33
Oakwood	-7	26
Rickardstown	20	21
Rossnagad	6	1
Summerhill	30	23
Tigroney	-4	1
Woodford	22	25

Uncorrected model	**Corrected model**
Hortland	
	
Kilduff	
	
Lackenrea 1	
	
Lackenrea 2	
	

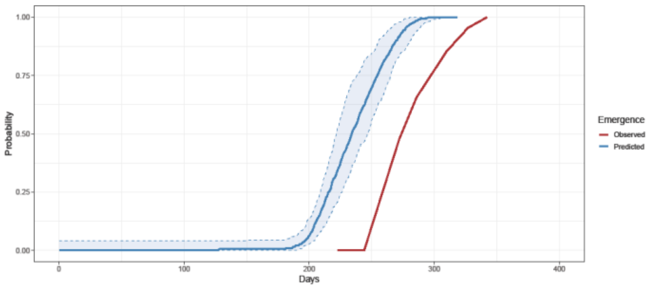
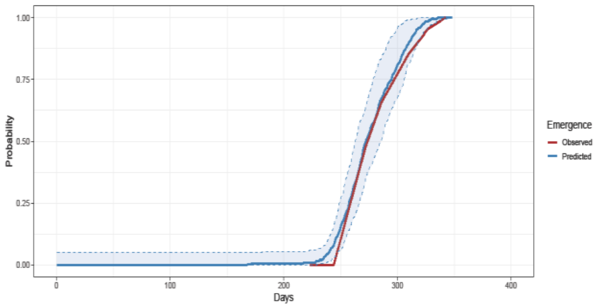
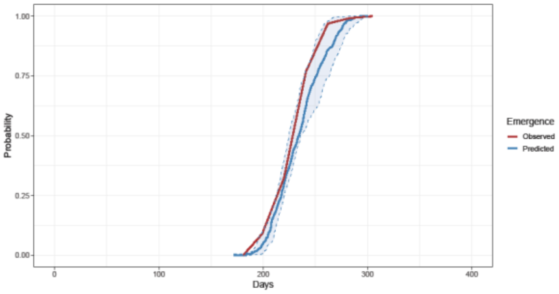
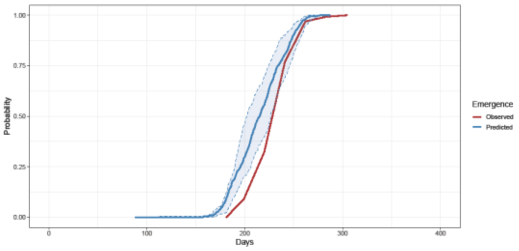
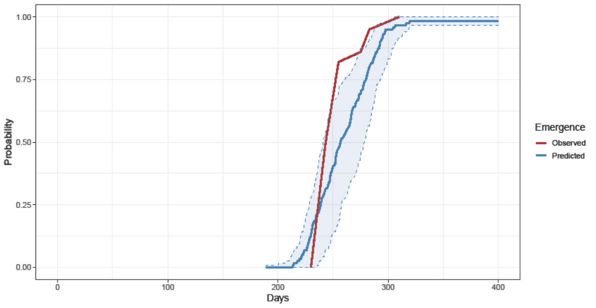
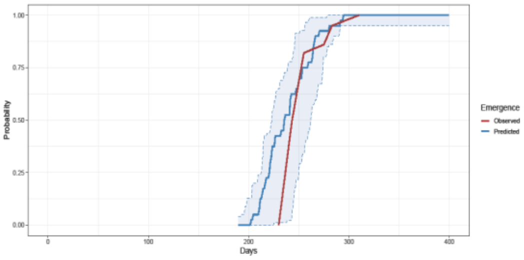
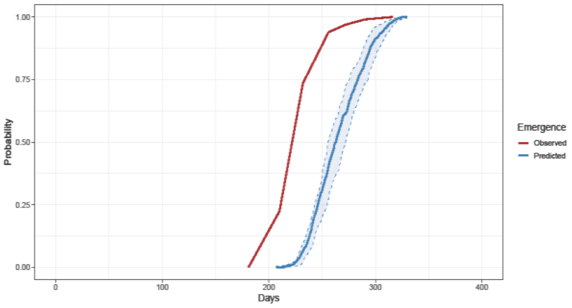
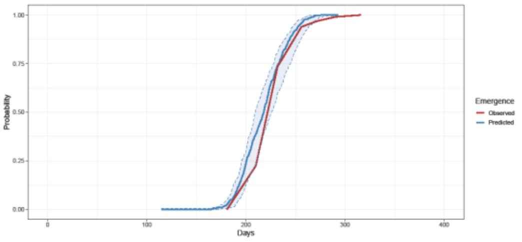
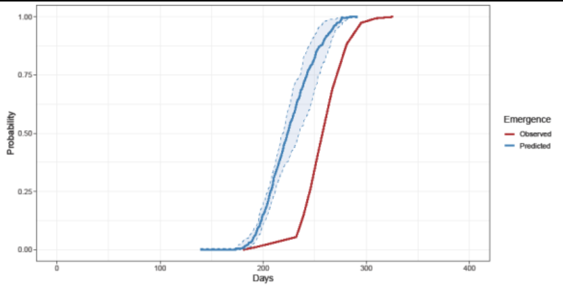
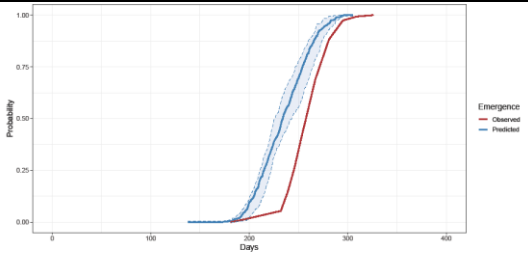
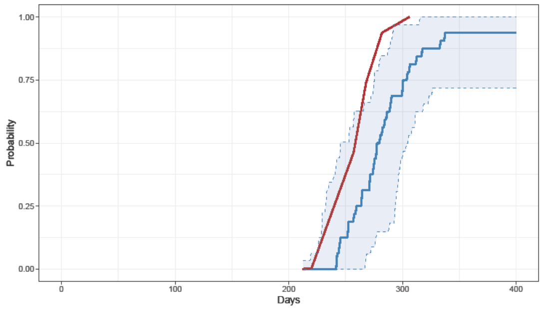
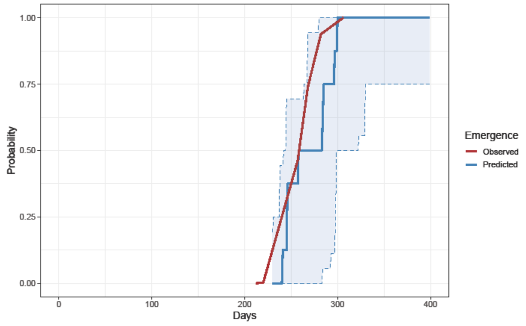
Gurtnapisha	
	

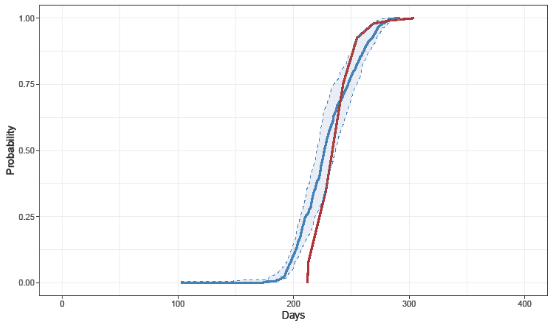
Figure A3: The observed and predicted simulation data (with 95% confidence intervals) for cumulative emergence of pine weevil at five trial sites, using weighted/co-located **met station temperature data**. Results for the uncorrected (left) and corrected (right) model.

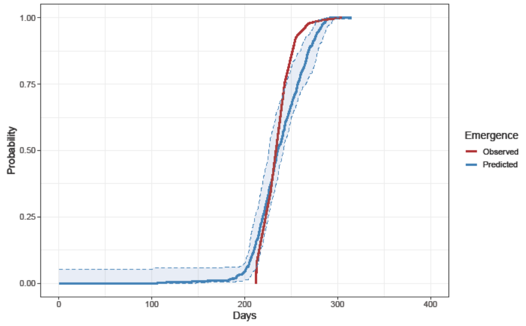
Uncorrected model	**Corrected model**
Hortland	
	
Ballinagee	
	
Kilduff	
	
Summerhill	
	
Glendine 1	



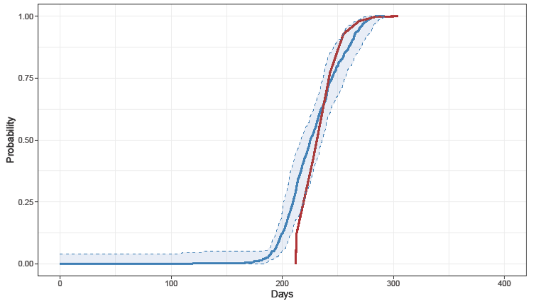


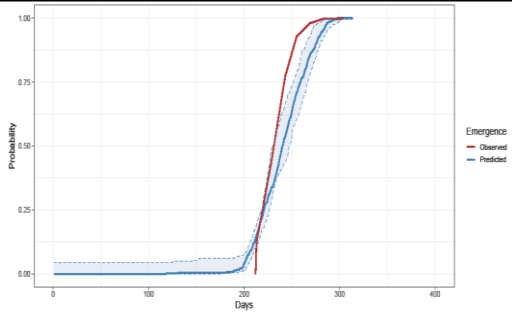
Lackenrea1



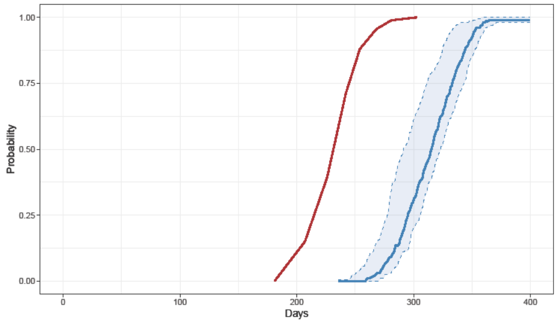


Lackenrea 2





Deerpark



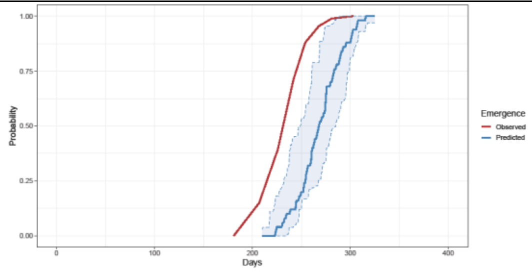


Figure A4: The observed and predicted simulation data (with 95% confidence intervals) for cumulative emergence of pine weevil at five trial sites, using **gridded temperature data (Walsh, 2012)**. Results for uncorrected (left) and corrected (right) model. Gortnpisha did not run.

Table A2: (Referencing Table 3.1 in main text*):* Site specific information of data inputs where trial sites were utilized (27) into the forecast modelling of `pineR`. Note that sites with one station are classified co-located sites.

Site Name	**Easting**	**Northing**	**Years (Trial years)**	**Co:**	**Weather stations used & Distances to trial sites***
Ballinagee	304739	202623	03-05'	Wicklow	*Ashford (22.8km),* *Genially (24.4km)*
Ballybrittas	258208	205198	01-03'	Wicklow	*Athy(17.7km),* *Durrow(32.7km),* *B. House(25.9km),* *Derrygreenagh(32.9km)*
Ballymacshaneboy	160232	118082	06-08'	Cork	*Mount Russell (2.08km)*
Ballyroan 1	248305	199495	99-01'	Laois	*Athy (21.5km),* *B. House (32.9km),* *Nealstown (29.2km),* *Derrygreenagh (30.4km)*
Ballyroan 2	248305	199495	00-02'	Laois	*Athy (21.5km),* *B. House (32.9km),* *Nealstown (29.2km),* *Derrygreenagh (30.4km)*
Cashelduff*	152794	298793	16-18'	Mayo	*Knock Airport (6.97km)*
Clonoghil*	224771	192091	14-16'	Laois	*Nealstown (5.58km)*
Cloondara*	206590	276631	13-15'	Longford	*Mount Dillon (5.46km)*
Deerpark	319829	213339	06-08'	Wicklow	*Ashford (17.2km),* *Glenealy (25.2km)*
Donadea	283467	232596	03-05'	Kildare	*B. House (26.2km),* *Derrygreenagh (34.1km)*
Doon	210178	231902	14-16'	Offaly	*Horseleap (19.2km),* *Gurteen (32.9km),* *Mullingar (40km)*
Glendine	228105	205254	12-14'	Offaly	*Nealstown (14.6km),* *Gurteen (29.1km),* *Durrow (30.8 km)*
Glendine trial 1	228105	205254	06-08'	Offaly	*Nealstown (14.6km),* *Gurteen (29.1km)*
Glendine trial 2	228105	205254	06-08'	Offaly	*Nealstown (14.6km),* *Gurteen (29.1km)*
Gurtnapisha*	229687	133751	14-16'	Tipperary	*Fethard (9.63km)*
Hortland	280071	235376	10-12'	Kildare	*Dunsany (19.3km),* *Derrygreenagh (30.2km),* *B. House (22.7km),* *Lullymore (13.2km)*
Kilduff*	245863	232411	10-12'	Westmeath	*Derrygreenagh (6.56 km)*
Killurney	226985	129529	13-15'	Tipperary	*Fethard (11.2km),* *Dungarvan (36.6km),* *Kilkenny-Gh. (35.3km)*

Knockaville	251079	249223	13-15’	Westmeath	*Mullingar (10.2km), Derrygreenagh (11.2km), B. House (18.3km), Horseleap (26.4km)*
Lackenrea 1	213159	101660	06-08’	Waterford	*Moorepark (30.8km), Dungarvan (11.4km), Ballincurrig (31.9km), Fethard (40.2km)*
Lackenrea 2	213159	101660	06-08’	Waterford	*Moorepark (30.8km), Dungarvan (11.4km), Ballincurrig (31.9km), Fethard (40.2km)*
Longfordpass	224385	160911	15-17’	Tipperary	*Fethard (20.6km), Durrow (23.5km), Kilkenny GH. (26.5km), Nealstown(32.8km).*
Oakwood	302744	201731	04-06’	Wicklow	*Ashford (24.3km), Glenealy (25.3km)*
Rickardstown	253419	263671	15-17’	Westmeath	*Mullingar (14.4km), Derrygreenagh (25.6km), B. House (31.8km)*
Rossnagad	243350	203335	11-13’	Laois	*Oakpark (37.5km), B. House (31.8km), Durrow (26.2km), Nealstown (25.0km)*
Summerhill*	284015	245951	09-11’	Meath	*Dunsany (8.2km)*
Tigroney*	320644	183264	14-16’	Tipperary	*Glenealy (6.48km)*
Woodford	170183	197082	16-18’	Galway	*Athenry (37.2km), Gurteen (29.2km)*

**Sites less than 10km were treated as co-located sites (one weather station used) within the weighted iteration of the* `pineR` *forecast model. They sites are highlighted in yellow above.*

Table A3. (Referencing Figure 5.1): Nearest Weather Station altitude and corresponding large pine weevil trial sites height of sites utilized in final modelling of `pineR`.

Trial Site Name	**Altitude**	**Nearest Met. Weather Station**	**Altitude**
Ballinagee	425m	Ashford	15m
Ballybrittas	68m	Athy	61m
Ballymacshaneboy*	323m	Mount Russell	195m
Ballyroan 1	4m	Nealstown	219m
Ballyroan 2	4m	Nealstown	219m
Tigroney*	246m	Glenealy	122m
Cashelduff*	215m	Knock Airport	201m
Clonoghil*	127m	Nealstown	219m
Cloondara*	41m	Mount Dillon	39m
Deerpark	260m	Ashford	15m
Donadea	64m	Dunsany	83m
Doon	71m	Horseleap	72m
Glendine	380m	Nealstown	219m
Glendine trial 1	380m	Nealstown	219m
Glendine trial 2	380m	Nealstown	219m
Gurtnapisha*	466m	Fethard	165m
Hortland	76m	Lullymore	85m
Kilduff*	93m	Derrygreenagh	94m
Killurney	371m	Fethard	165m
Knockaville	95m	Derrygreenagh	94m
Lackenrea 1	170m	Dungarvan	18m
Lackenrea 2	175m	Dungarvan	18m
Longfordpass	130m	Fethard (PH)	165m
Oakwood	250m	Glenealy	122m.
Rickardstown	89m	Mullingar	101m
Rossnagad	94m	Nealstown	219m
Summerhill*	90m	Dunsany	83m

*Co-located sites (one weather station used in model) are marked above.

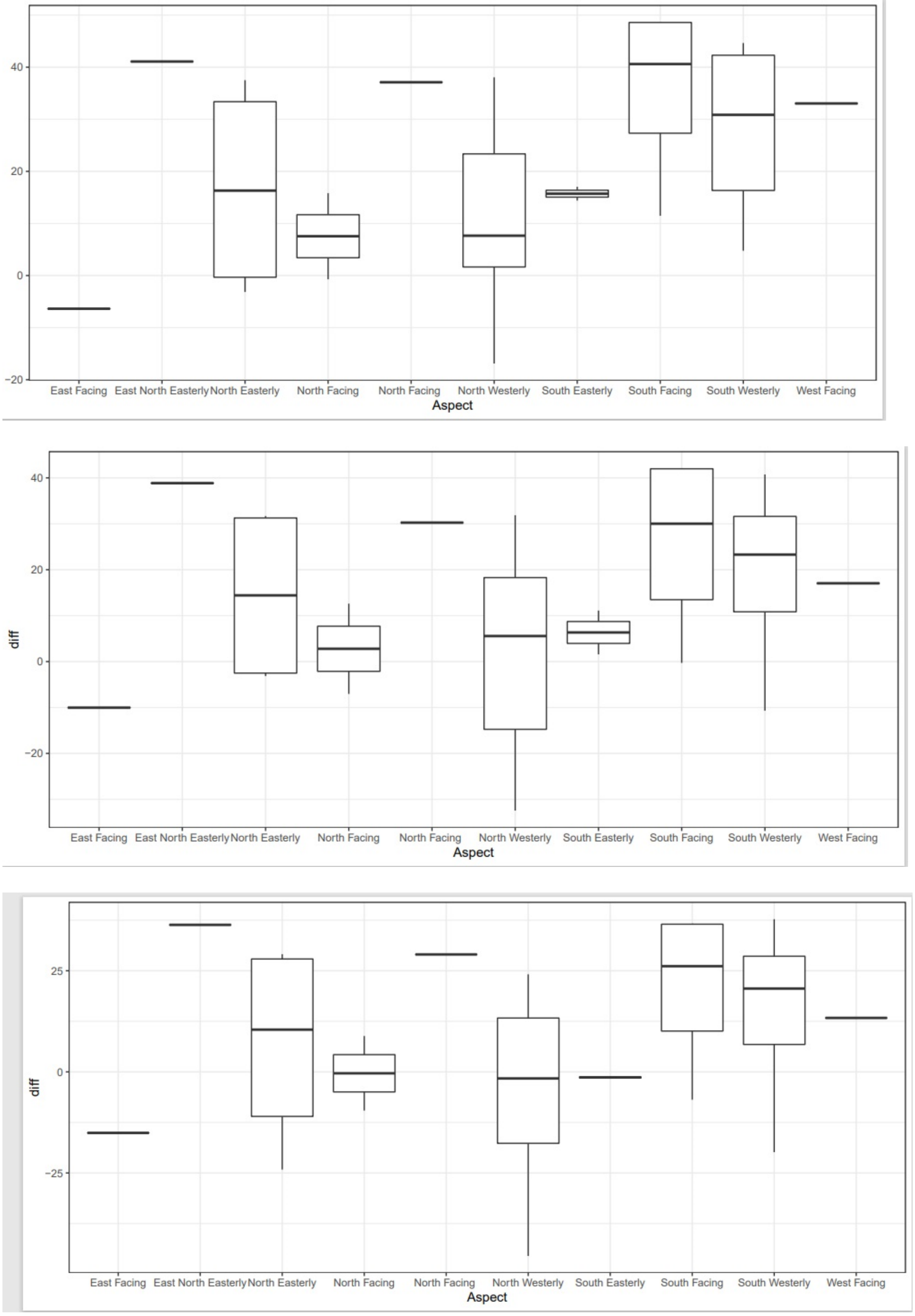


Figure A5-A7: Boxplots displaying influence of trial sites aspect on `pineR` utilising data at all 27 trial sites prior to correction been implemented or applied in original forecast model. Boxplots are covering influence of aspect at each depth inputted in original model in order of 10cm, 20cm, 30cm.