The temporal and spatial ecology of *Brassicogethes (Meligethes) aeneus*.

Doctor of Philosophy

School of Agriculture, Policy and Development

Chris R. Shortall

January 2021

Declaration:

I confirm that this is my own work and the use of all material from other sources has been properly and fully acknowledged.

Chris R. Shortall

Slip End, Bedfordshire, January 2021

Abstract

Pollen beetles (*Brassicogethes (Meligethes) aeneus*) are a common and pernicious pest of oilseed rape. This study investigated the migration ecology of *B. aeneus* using suction-trap data from the longest-running terrestrial insect dataset in the world alongside field-collected data from a five-year sampling campaign. In addition, next-generation sequencing techniques and flight-mills were used to investigate the flight behaviour and migratory potential of *B. aeneus.* Populations have increased in abundance by ~162% in the UK since the 1980s alongside, but not correlated with, an increase in oilseed rape. The abundance of spring migrants in spring is related to the abundance of beetles caught in suction-traps the previous summer, suggesting that it may be possible to forecast the scale of the spring migration into crops. The number of *B. aeneus* in daily suction-trap samples is influenced by a small range of environmental variables, time of year and the number of beetles trapped in the previous summer. A selection of statistical approaches (GAM, zero-inflated GLM, random forest and an artificial neural network) were used to investigate daily suction-trap samples, however, none were able to predict the eruptive dynamics that quickly lead to very high counts. The spatial ecology of *B. aeneus* showed a strong positive relationship between suction-trap and field-caught data and suction-trap data showed spatial synchrony of migrating beetles up to ~150 km. However, no spatial synchrony was found in the field-caught data indicating that at the farm scale the beetle may be locally irruptive, and infestations of the crop are difficult to predict. Studies on the dispersal ability of *B. aeneus* in the UK proved inconclusive, but indicated that there is a single, homogeneous, UK population with a proportion migrating long distances within the UK. In addition, studies on adult *B. aeneus* emergence in the field and collated life cycle parameters for *B. aeneus* are presented as appendices. Together, these findings demonstrate the potential to use suction-traps to forecast the timing and abundance of *B. aeneus* spring migration and provide growers with timely warnings for crop inspections and control.

Acknowledgements

This research has taken a long time. When this project started the world was sane, Brexit and Covid were not even words and Donald Trump was no more than another TV personality. To have been able to plug on through these unprecedented times meant that I needed a strong support network and I have many thanks to make.

Firstly, to my supervisors, James, Sam, Alice and Julian, thank you for helping me through the rough patches, prodding me through the slow bits and teaching me all I needed to know. Thank you for being cheerleaders and task masters when you needed to be.

To my colleagues in the Insect Survey: Tracey, Mike, Lynda, Alex, Gemma, Alex, Dan and Aislinn, thank you for your unwavering support throughout the time. Thank you for volunteering your time to count beetles in addition to all the other tasks you had to perform. Thank you for sharing the burdens in the light-trap network and suction-trap archive to allow me to concentrate on getting this work done. And thank you for providing the space to vent my frustrations and for distracting me from my problems when needed.

To my colleagues at Rothamsted, thank you for your help when needed, to Ramiro, Dion and Fernando thank you for guiding me through the perils and pitfalls of molecular work. To Steve and Rachel thank you for allowing me into your lab and your generosity in time and patience in showing me the ropes and running samples through the bioanalyzer for me. To Suzanne, Yoann & Kirsty for your statistical help and support, I am indebted to your knowledge and patience. To Martin and Todd, thank you for your assistance in the field. To Donna for your unwavering support to me and all the other students, you are a credit to the Institute. To Liesel, though you didn't know it you twice said exactly the right thing at the right time to set me back on course when I wavered.

To Rothamsted Cricket Club, St John Fisher Vets and The Bush CC for providing me with the opportunity to hit a ball hard, run around a lot and release some pressure.

To The Undesirables for keeping me grounded and reminding me there is a world outside academia with its own issues and inspirations.

To my parents for all your support and assistance, not just during this project, but all my life – I wouldn’t be where I was without you.

Last and very much not least, to Alison, Charlotte and Rory for keeping me sane and happy, I will forever be thankful for your patience over these last six years and for a level of support so far beyond what could be expected. I love you all, and now it is time for me to repay you with my time and energies – I owe you so much

Contents

List of Figures

List of Tables

Chapter 1.

Introduction to oilseed rape and pollen beetles

1.1 Oilseed rape

1.1.1 History of oilseed rape as a crop plant

The early history of oilseed rape (*Brassica napus* L., OSR) use in the UK and the rest of Europe is difficult to interpret due to a lack of consistent nomenclature in the literature, causing confusion with other similar crops in older manuscripts, particularly turnip rape *B. rapa* L. (Bunting 1986). Probably arising as a hybrid between turnip rape and cabbage (*B. oleracea* L.) in the Middle Ages (Allender & King 2010), OSR was almost certainly introduced to England by Dutch engineers working on fenland drainage projects in the late 16th or early 17th century (Fussell 1955). Although the oil produced from the seeds was too high in glucosinolates to be palatable for food, rapeseed production, especially in the East of England, increased rapidly through the 17th century and rapeseed oil became the preferred product for oil lamps, with the residual oil used as lubricant (Lamb 1989) or cattle feed (Bunting 1986). Through the 18th century OSR cultivation gradually extended into other areas of the country but by the 19th century processing plants were being set up at ports as the emphasis shifted to imported seed (Bunting 1986). By the mid-19th century there was still some OSR growing in the fens of Eastern England, but its use was more as a forage crop for sheep. Mineral oil had largely replaced organic oils for lighting and lubrication and the crop was effectively absent from British agriculture for that reason (Bunting 1986).

Throughout the early 20th century OSR remained unused in the UK; there were small amounts imported from China and India for non-edible uses, particularly as an engine lubricant, but the levels were so low as to remain insignificant (Bunting 1986). In Europe there was a brief period of increased production during the war years (Appelqvist *et al.* 1973) but it was not until the middle of the century that production began to increase in Europe, alongside a general surge in production worldwide (Appelqvist *et al.* 1973, Bunting 1986). In Britain the crop remained a minor element as a break crop in intensive wheat cultivation (Appelqvist *et al.* 1973, Almond *et al.* 1986), with production levels per unit area much lower than the rest of Europe until the mid-1970s (Bunting, 1984).

Research in the 1960s - 70s, particularly in Canada, to reduce glucosinolate content resulted in varieties (for example canola) that were suitable for use in food production and the economic value of the crop increased significantly (Downey & Craig 1964, Downey 1987, Wittkop *et al*. 2009, Freidt *et al*. 2018). These new varieties led to a rapid rise in the area under OSR cultivation in the UK concomitant with its increased economic value and relaxation of regulatory pressures on cropping (Champion *et al*. 2003, Hillocks 2012).

From 1984 to 1994 production in the UK declined but by 2004 the EU was the largest producer of rapeseed in the world, with the UK the third largest producer within the EU (Newcastle University, 2006). After a decade of fluctuation, since 2004 yields have improved (Knight *et al*. 2012) alongside an increase in area under OSR crop (Gehringer *et al*. 2007), with 2012 seeing a record 737,000 hectares for harvest in Great Britain alone (HGCA 2013). OSR is now grown in the UK for the oil contained within the seed which is used for cooking oils and margarines and as vegetable oil in processed foods (Freidt & Snowdon 2009). The meal left after oil extraction is utilised as a high protein animal feed (Lamb, 1989). Increasingly, the potential for OSR and other oleaginous crops are being investigated for exploitation within the bioeconomy, for example lubricants, plastics, soaps and biofuel (Dumont & Narine 2008, Wu *et al*. 2009, Manamperi *et al*. 2010, Zainal *et al*. 2018).

1.1.2 Current agronomy

Oilseed rape is usually grown as a break crop in cereal rotation (Christen & Sieling 1995), which has three primary benefits: an effective break into higher yielding first cereal crops, control of grass weeds and an early drilling window which spreads farm workload (HGCA 2013). Initially oilseed rape was grown in one in five rotations with cereals, however economic pressure has led to shorter rotations and one in three, or even one in two rotations have become more common (HGCA 2013). Variety selection is primarily governed by gross output, with other qualities considered being: stem factors (stiffness, shortness, resistance to lodging), maturation factors (earliness of flowering and maturity), oil and glucosinolate content, and disease resistance factors (particularly light leaf spot and phoma stem canker resistance) (HGCA 2013). There are four classes of variety: conventional open-pollinated which are wind and/or self-pollinating and can be grown from farm-saved seed without paying breeders royalties, Clearfield ® herbicide tolerant hybrids for use with specific herbicides, semi-dwarf hybrids which produce short, stiff-stemmed plants which are easy to harvest but have lower yields and restored hybrids created by crossing male-sterile ‘female’ plants with pollen producing ‘male’ plants. Hybrid cultivar seed cannot be farm-saved without royalty payments but are potentially more adaptable to a wide range of growing conditions (Thompson & Hughes 1986, HGCA 2013). In addition there are varieties for specific markets; High eruic acid rape (HEAR) varieties are for use in industrial processes such as inks, lubrication etc. and High oleic, low linoleic (HOLL) rape varieties which produce oils that are low in trans fatty acids and saturated fats and are used as cooking oils (Thompson & Hughes 1986, HGCA 2013). Care must be taken to avoid contamination between HOLL and other varieties that are unpalatable, or unfit for consumption, particularly HEAR to the extent that land used for HEAR cultivation should never be used for HOLL crops (HGCA 2013).

Yield is measured by 1000 grain weight, with oil content an important factor. Seed number, usually 15-20 seeds per pod in OSR (Diepenbrock 2000) is determined by the amount of photosynthesis carried out by the crop during the period 19-25 days after flowering (Habekotté 1997a, Weightman et al. 2010, Bennett et al. 2017); for best yields the optimum density of pods for efficient photosynthesis 6,000-8,000 pods per m^2 (Berry & Spink 2006, Weightman *et al*. 2010), a lesser density indicates the canopy size is too small to trap all incident radiation, whereas a greater density will likely mean the flowering layer is preventing sunlight reaching the photosynthetic tissues (Berry & Spink 2006). Seed weight is determined by the length of the seed filling period and amount of photosynthesis during that time (Habekotté 1997a, b, c, Weightman *et al*. 2010); factors that can curtail seed filling are shallow rooting, drought, disease, early desiccation and lodging causing sub-optimal canopy structure (HGCA, 2013). There is evidence that increased winter temperatures cause yield loss, although the mechanism is still unclear (Brown *et al*. 2019). Pest and disease pressure can also reduce the number and quality of seeds per pod and the number of pods per plant (Williams 2010).

Inputs into an oilseed rape crop are relatively high (Muthmann 2007) and include: nutrient provision, primarily nitrogen and sulphur but also phosphorous, potassium and magnesium for long-term soil maintenance and the micronutrients boron, manganese and molybdenum; fungicides, primarily to control phoma leaf spot, light leaf spot, *Sclerotina* stem rot, powdery mildew and grey mould; molluscicides to control slugs; herbicides including graminicides to target black grass which is an increasing problem in cereal crops, pre-emergent herbicides for broad-leaf weed control with some post-emergent herbicides for specific weed problems (e.g. cleavers and crane’s-bill); desiccant application at harvest; insecticides to control specific insect pests, the most important being the aphids *Myzus persicae* (Sulzer) and *Brevicoryne brassicae* (L.), the flies *Dasineura brassicae* (Winnertz) and *Delia radicum* (L.), and the beetles *Psylliodes chrysocephala* L., *Ceutorhynchus obstrictus* (Marsham) and the pollen beetle *Brassicogethes aeneus* (Fab.); the most numerous insect found on OSR and subject of this thesis. Fig. 1.1 shows the timing of the major interventions through a winter OSR growth season.

Due to concerns expressed about the potential of neonicotinoids to harm bees and wild pollinators (Gill *et al*. 2012, Whitehorn *et al*. 2012, Henry *et al*. 2012), though disputed by some (Cresswell & Thompson 2012, Carreck & Ratnieks 2014), neonicotinoid seed dressings – used in the UK to protect OSR from aphids and flea beetles (Scott & Bilsborrow 2019) - were subjected to an EU-wide moratorium on flowering crops from December 2013 (OJEU 2013, Scott & Bilsborrow 2019), leading in turn to a reduction in area of OSR in most countries (Dewar 2017, Carreck 2017, Scott & Bilsborrow 2019), with the crop being replaced by barley, field beans, peas or sugar beet in many situations in the UK (de la Pasture 2016, Stephens *pers. comm*.).

1.2 Pollen Beetles

1.2.1 Taxonomy & nomenclature

Pollen beetles, or blossom beetles, are a group of small (1-5mm) species within the subfamily Meligethinae (Coleoptera: Nitidulidae). The subfamily has a global distribution, absent only from the Neotropical and Antarctic regions with the approximately 700 known species being anthophagous (i.e. flower-feeding), of which some species are monophagous, or narrowly oligophagous as larvae (Kirk-Spriggs 1996, Audisio *et al*. 2009). The species in the genus *Meligethes s.l.* Stephens are characterised by the presence of semi-circular impressed lines on the base of the pygidium (Fig. 1.2a) and the outer edge of the anterior tibiae are often toothed (Fig. 1.2b) (Kirk-Spriggs 1996, Audisio *et al*. 2009); the size and arrangement of these tibial teeth is often an aid to

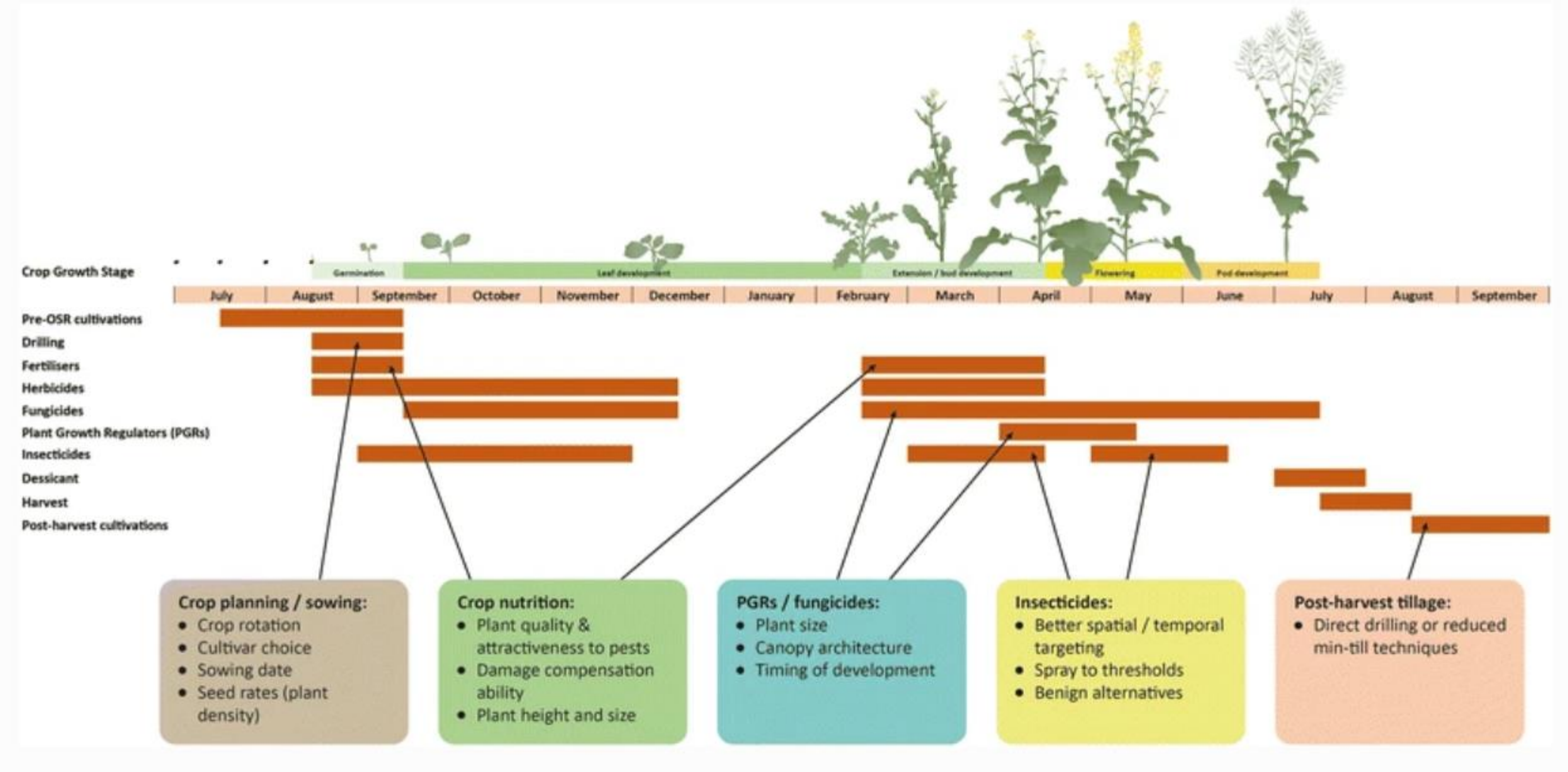


Fig. 1.1: Oilseed rape crop management practices in relation to crop growth stage, from Skellern & Cook (2017).

species identification, however there is a great deal of inter- and intra-specific variation in these and other identification characters, making identification of the group somewhat difficult (Kirk-Spriggs 1996). A recent review of the Meligethinae by Audisio *et al*. (2009) created a suite of new genera within the subfamily and reassigned many species, including *Brassicogethes (Meligethes) aeneus* (Fab.) and *B. (Meligethes) viridescens* (Fab.), both species specialising on

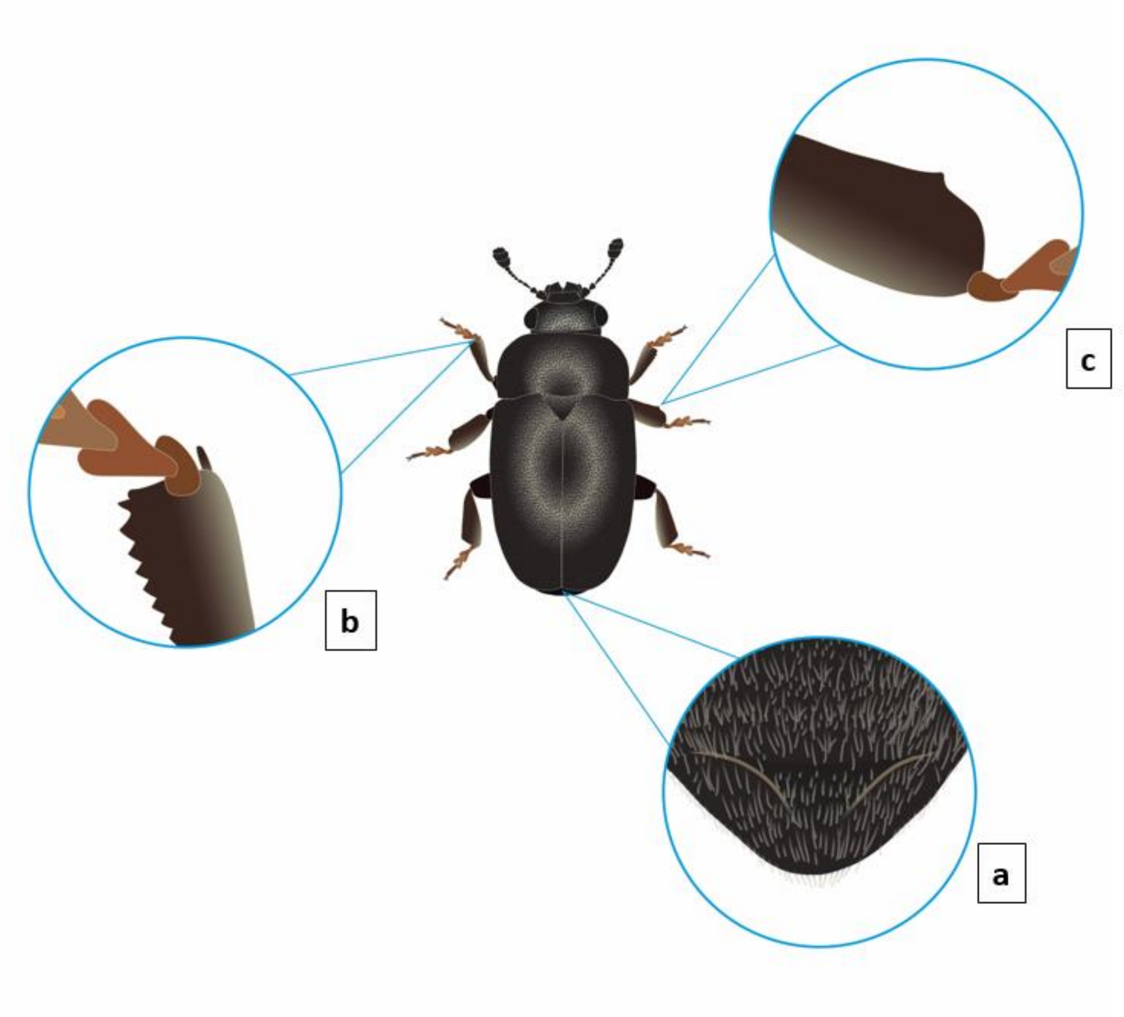


Fig. 1.2: Diagram of the pollen beetle (*Brassicogethes aeneus*), showing the distinctive semi-circular depressions on the pygidium (a), the fore-tibia (b) and the middle femur of *Brassicogethes viridescens* (c) showing the diagnostic projection on the lower edge of the middle femur that is absent from *Brassicogethes aeneus*.

brassicas as a larval food plant (Kirk-Spriggs 1996, Audisio *et al*. 2009), the latter having a projection on the lower edge of the middle femur (Fig. 1.2c) and slightly larger & more diffuse puncturation (Hoebeke & Wheeler 1996, Kirk-Spriggs 1996).

1.2.2 Pest Status

Although most species are of no economic importance, some species of *Meligethes s.l.* are considered pests; in Europe *M. flavimanus* Stephens attacks cultivated roses, though only rarely (Fowler 1885, Audisio *et al*. 2014) and *Fabogethes (Meligethes) nigrescens* (Stephens) feeds on decorative sweet pea flowers (Kirk-Spriggs 1996) but is known as a pest on cultivated clover in North America (Hatch 1957). In Europe only *B. aeneus* and, to a lesser extent, *B. viridescens* are economically important crop pests (Williams & Free 1978, Valantin-Morison *et al*. 2007, Valantin-Morison & Pinochet 2010, Williams 2010). Both species feed on cruciferous crops and can cause serious yield loss in agricultural crops – mainly OSR, through adult feeding on buds, causing them to abscise (Free & Williams 1978, Winfield 1992). In horticultural crops such as cauliflower and broccoli the feeding of adults in the flower heads can cause spoilage making them unsellable (Finch *et al*. 1990). The larvae of both species also feed on pollen and nectaries (Cook *et al*. 2004a, Seimandi Corda *et al*. 2018) but rarely cause the significant yield loss associated with adult feeding (Free & Williams 1978).

Brassicogethes viridescens has a later phenology and usually arrives in crops once they are past their susceptible stage, and at lower abundance meaning that it is less economically important than *B. aeneus* in winter OSR (Alford et al. 2003, Hiiesaar et al. 2003). The species can occasionally be a significant pest in spring OSR when temperatures are high enough for them to occur at damaging levels of abundance (Finch et al. 1990, Alford et al. 2003). *Brassicogethes viridescens* was introduced to North America in the mid-20th century (Hoebeke & Wheeler 1996, Mason *et al*. 2003) and, though currently restricted to eastern Canada and Maine (Hoebeke & Wheeler 1996, Mason *et al*. 2003), could prove to become a key pest of North American canola crops (Olfert & Weiss 2006, Dosdall & Mason 2010). However, since most of the work on pollen beetles in Europe has been based on *B. aeneus*, this species will be the focus of this research.

In Europe the abundance and early emergence of *B. aeneus* means that it is a major pest of both winter and spring OSR (Finch *et al*. 1990, Alford *et al*. 2003, Hansen 2004). Whilst females chew oviposition holes in the base of developing buds (Ekbom & Borg 1996), the primary damage is through adults chewing holes in unopened flowers to access the pollen, causing them to abscise (Fig. 1.3 inset) and resulting in loss of yield (Nilsson 1987, Cook *et al*. 2004a, Ferguson *et al*. 2015). This rarely causes significant impact on winter OSR yields as the crop is usually well into flower by the time the beetles emerge from hibernation and migrate in any significant numbers into the crop (Williams 2010). However, late-flowering varieties of winter OSR or those with delayed development due to later sowing, poor emergence conditions or pigeon damage can be vulnerable to *B. aeneus* attack (Hiiesaar *et al*. 2003) especially when population abundance is high and when warm conditions trigger large mass migrations into the crop within a short period of time (Zlof 2008). Spring OSR can, however, be seriously affected by *B. aeneus* attack with yield losses of up to 80% reported in some years (Hansen 2004), in large part due to the large population abundance of all emerged overwintered adults as well as some new generation beetles. In addition, female *B. aeneus* mature their eggs continuously through their life (Ekbom & Ferdinand 2003) and thus are able to move onto spring OSR to feed and reproduce when it is at the vulnerable green bud stage (Nilsson 1987, Ekbom 2010). Fig. 1.3 details the timing of the different life stages of *B. aeneus* alongside the development of the oilseed rape crops in the UK, experimentally and observationally derived parameters for each stage are collated from the literature and presented in Appendix D.

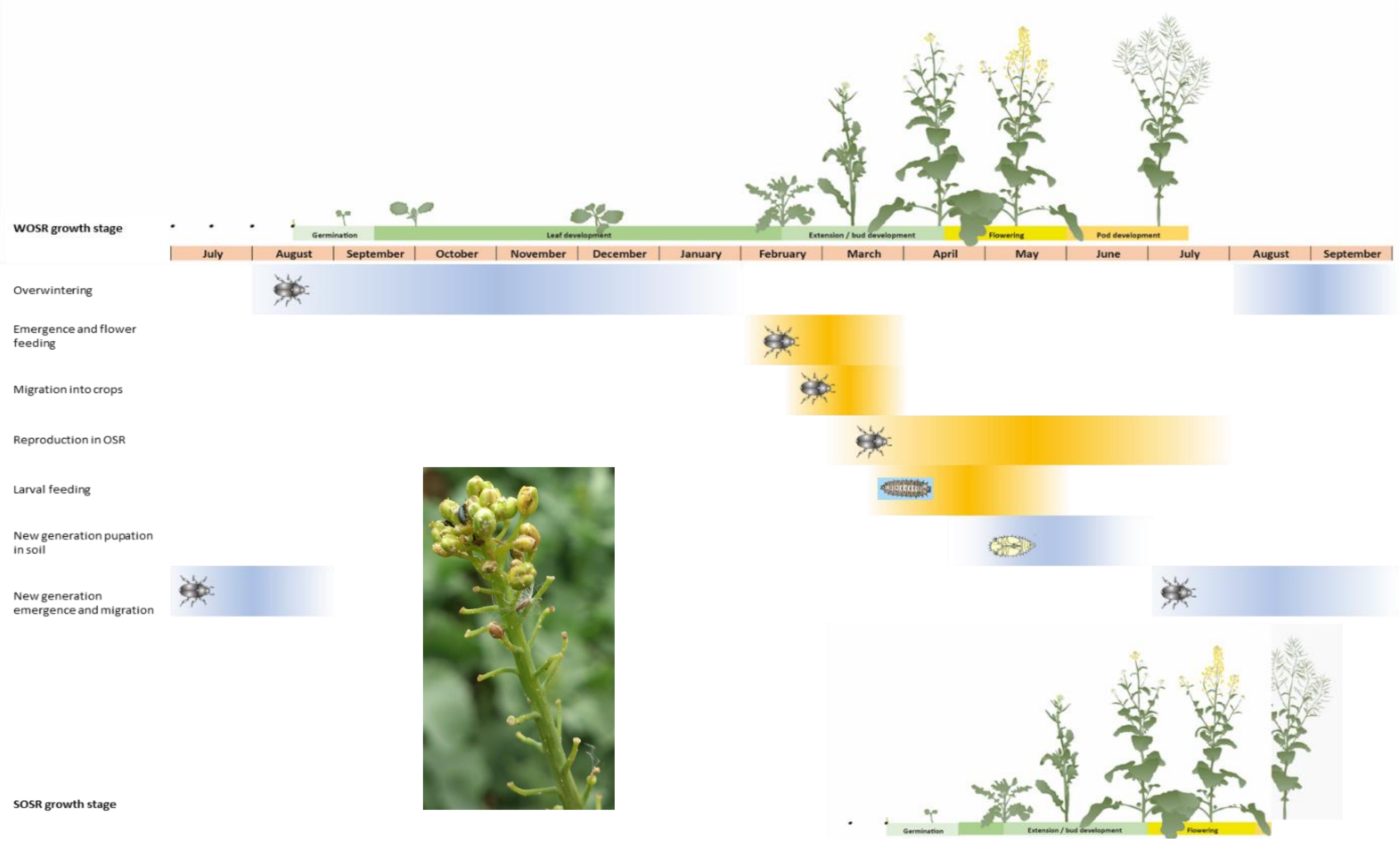


Fig. 1.3: Life cycle of the pollen beetle (*Brassicogethes aeneus*) alongside the crop development stage of winter (top) and spring oilseed rape (bottom). Orange coloured bars indicate life stages where the beetle can be considered a pest. The stronger colours in the life stage bar indicates the peak period. The inset photo shows typical damage caused by the beetle, podless stalks caused by abscission of damaged buds. (Photo courtesy of Sam Cook).

Pre-flowering infestation of an OSR crop often results in the appearance of podless stalks due to the loss of buds (Williams & Free 1979). However, brassicas naturally abort up to 50-60% of their flowers and buds, therefore insect damage can be compensated for by retaining those that would otherwise have been aborted (Williams & Free 1979, Tatchell 1983, Lamb 1989). It is thought that OSR can withstand at least 6-8 *B. aeneus* per plant at the 'separated green bud stage' without yield loss, even though there is visible damage to the crop (Lerin 1988). However, higher numbers of *B. aeneus* can lead to yield loss (Zlof 2008).

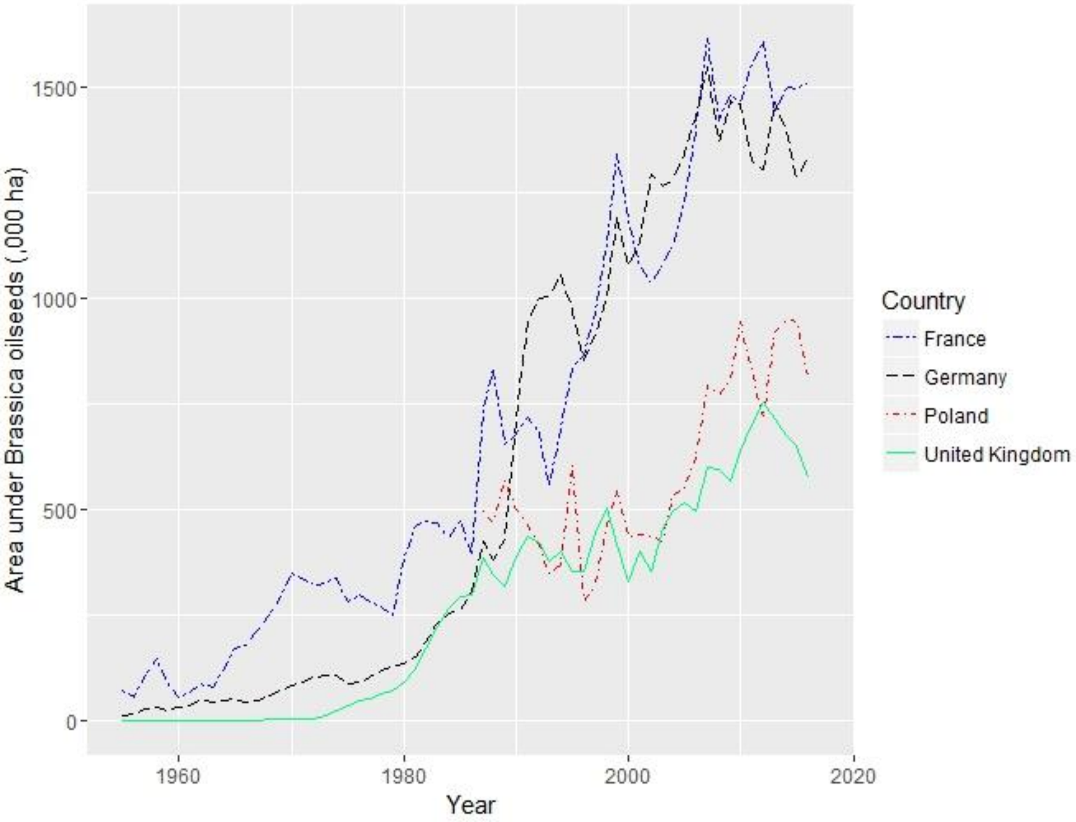


Fig. 1.4: Area (1000's ha) of brassica oilseeds *(B. napus* and *B. rapa*) grown in the top four producing countries in Europe (1955 -2016, Poland: 1987-2016): data from Eurostat.

Brassicogethes aeneus has been shown to be able to rapidly shift from native hosts to OSR when the crop is introduced to an area (Hokkanen 2000) and shows a significant increase in reproductive success on OSR compared to populations on the natural host (Hokkanen 2000). This adaptive advantage - i.e. being able to rapidly exploit a new host - appears key to the success of *B. aeneus* as a pest and there is some indication that conventional plant management (i.e. synthetic insecticide use) may enhance the adaption of this species (Hokkanen 2000). This ability to shift to OSR allowed it to exploit the increase in acreage that occurred in Europe in the second half of the 20th century – particularly in the four main countries where OSR is grown (Fig. 1.4) – and become an abundant pest throughout the region.

1.2.3 Biology, ecology and life cycle of Brassicogethes aeneus

1.2.3.1 Basic life cycle

The basic aspects of the life cycle of *B. aeneus* are well-known (Fig. 1.3, Appendix D). Adults overwinter in hibernation in the leaf litter and top layers of soil in woodlands and hedges (Hokkanen 1993, Alford *et al*. 2003, Rusch *et al*. 2012, Juhel *et al.* 2017) before emerging in the spring and migrating either directly to their larval host plants or via feeding stops on the pollen of other flowering plants (Williams & Free, 1978, Ouvrard *et al*. 2016, Juhel *et al.* 2017). The threshold temperature for emergence is still an open question, however the lowest temperature recorded for a solitary flight is 5.9°C in a laboratory environment (Ferguson *et al*. 2015) and 10.2°C in the field (Láska & Kocourek 1991); gregarious migrations (Cooter 1977, Kenward 1984) have been predicted for a mean air temperature for migration of 8°C in a phenological model for Luxembourg (Junk *et al.* 2016) and noted in-field at temperatures in the range 12.3°C - 15°C (Taimr *et al*. 1967, Tulisalo & Tuomo 1986, Nilsson 1988a, Finch *et al.* 1990, Láska & Kocourek 1991, Šedivý & Kocourek 1994, Ferguson *et al.* 2013, 2015, Skellern *et al*. 2017). Flights are predominantly diurnal (Mauchline *et al.* 2017), with Lewis & Taylor (1965) establishing a peak flight time of 12:44 GMT.

The adults of *B. aeneus* feed on pollen from many plant families (Fritzsche 1957, Free & Williams 1978, Williams & Free 1978, Williams 2010, Ouvrard *et al*. 2016) from which they obtain nutrients such as proteins, amino acids, lipids, sterols, starch, vitamins and minerals (Roulston & Cane 2000) and can be abundant, comprising over 25% of flower visitors in one survey (Baldock *et al*. 2015). However, as the larvae are monophagous, they oviposit exclusively on brassicas (Free & Williams 1978, Williams & Free 1978, Charpentier 1985, Nilsson 1988c, Ekbom & Borg 1996, Bartlet *et al*. 2004, Cook *et al*. 2004a). *Brassica* spp. L. are particularly favoured for oviposition (Veromann *et al*. 2012, Kovács *et al*., 2013), however, other studies find *Sinapis* spp L. and *Raphanus* spp. L. are equally, or more attractive (Kaasik *et al*. 2014a, b).

On reaching the larval host plant, females lay eggs inside the developing flower buds (Free & Williams 1978, Charpentier 1985) preferring buds in the 2-3mm size range (Ekbom & Borg 1996, Ferguson *et al*. 2015). The eggs are 0.8mm-3mm long (Ekbom & Popov 2004; Ekbom & Borg 1996) with egg size being smaller on lower-quality hosts (Ekbom & Popov 2004). Each female can lay more than 200 eggs over a two-month period (Hopkins & Ekbom 1996), laying up to 3 eggs per bud (Ekbom 1998) with clutch size being larger on more suitable hosts (Hopkins & Ekbom 1999, Ekbom & Popov 2004). Most eggs are laid in the morning and early afternoon (Ekbom & Ferdinand 2003) with oviposition behaviour generally positively related to temperature (Ferguson *et al*. 2015). Egg development usually takes around 4-7 days (Cook 2000, Cook *et al*. 2004a, 2006b) and there are two larval instars (Cook 2000), feeding mainly on pollen but all parts of the flower may be destroyed (Giamoustaris & Mithen 1996). The first instar usually feeds on the pollen of the developing anthers inside the closed flower bud. In warmer weather buds sometimes open before the larvae hatch causing the eggs to perish (Nilsson 1987). Upon budburst the larvae usually moult to the 2nd instar and larvae feed on the pollen from open flowers, moving up the plant (Free & Williams 1978). Development time is between 9-13 days under laboratory conditions (Osborne 1965, Cook *et al*. 2004a, Seimandi Corda *et al*. 2018), on reaching maturity the larvae drop to the soil to pupate within an earthen cell. Pupation takes a few weeks before the adults emerge and feed on flowers of several late-summer flowering plants, including spring oilseed rape, before seeking shelter over winter (Hokkanen 1993, Valantin-Morison *et al*. 2007, Juhel *et al*. 2017).

1.2.3.2 Natural enemies

As small beetles with an apparent lack of predator defence such as toxins, spines or startle colouration it is expected that *B. aeneus* is a target for most generalist predators, both invertebrate and vertebrate (Holling 1961, Buckner 1966, Taylor 2013). The species is known to form a significant part of the diet of the common swift (*Apus apus* (L.)) (Lack & Owen 1955) and they are recorded as prey of a range of invertebrates (reviewed by Skellern & Cook, (2018)) including predators such as spiders (Wolff & Krausse 1926, Osborne 1960, Harenberg 1997, Cassel-Lundhagen *et al*. 2009), beetles (Friedrichs 1921, Osborne 1960, Williams *et al*. 2010) and lacewings (Börner & Blunck 1920, Osborne 1960). These predators have been demonstrated to have a depressant effect on *B. aeneus* abundance in OSR fields (Büchi 2002, Büchs & Alford 2003, Hokkanen 2004, Frank *et al*. 2010, Williams *et al*. 2010) with losses to predators of 50-80% reported (Büchs & Nuss 2000, Büchi 2002), although some studies suggest much lower predation rates (Hokkanen *et al*. 1988). In addition, there is a suite of parasitoids (parasitic wasps that lay eggs on or in the host, usually killing it during development) known to attack *B. aeneus* larvae (Osborne 1960, Ferguson *et al*. 2010, Ulber *et al*. 2010, Rusch *et al*. 2011, 2013, Vinatier *et al*. 2012). The four key, economically important species are: *Phradis interstitialis* (Thomas); *P. morionellus* (Holmgren);

Tersicholus heterocerus Thompson; and *Diospilus capito* (Nees) (Ferguson *et al*. 2010). Other recorded parasitoids are: *Aneuclis incidens* (Thompson); *Blacus nigricornis* Haeselbarth; *Eubazus sigalphoides* (Marshall); *Cerchysiella planiscutellum* (Mercet); and *Brachyserphus parvulus* (Nees) (Osborne 1960, Ulber *et al*. 2010). Reported rates of parasitism of *B. aeneus* in the field vary widely between studies from 0% to more than 90% (Nilsson & Andreasson 1987, Nielsen & Axelsen 1988a, Hokkanen 2000, Büchi 2002, Williams 2006, Thies *et al*. 2008, Tölle 2014, Riggi *et al*. 2017, Seimandi Corda 2018).

1.2.3.3 Migration

As mentioned above (Section 1.2.3.1), aside from localised movements between feeding sites, *B. aeneus* migrates twice a year, between overwintering sites and reproductive sites in spring and between feeding and overwintering sites in late summer (Williams & Free, 1978, Ouvrard *et al*. 2016, Juhel *et al.* 2017). How far individuals are able to migrate is a matter of conjecture; Taimr *et al.* (1967) used radioactive tracers to establish that they can travel at least 13.5 km whilst more recently, Juhel *et al*. (2017) used geomatics to model distances from woodland edge to infested fields and estimated mean dispersal at 1.2 km. *Brassicogethes aeneus has* been recorded by high altitude (200m) nets and vertical-looking radar (Mauchline 2003, Mauchline *et al.* 2017) indicating that their potential migration distance is much further. It should be noted that *B. aeneus* has not been recorded in offshore traps (Hardy & Milne 1937, 1938, Glick 1939, Gressitt *et al*. 1962, Hardy & Cheng 1986) which may indicate that long-distance migrations are uncommon.

1.2.3.4 Host plant location

The primary visual cue for *B. aeneus* is colour; they are known to be differentially attracted to the colour yellow (Wassman 1926, Láska *et al.* 1986, Buechi 1990, Evans & Allen-Williams 1994, Giomoustaris & Mithen 1996, Döring *et al*. 2012, Cook *et al*. 2013b). Traps coloured yellow attract more *B. aeneus* than other colours (Blight & Smart 1999, Williams *et al*. 2003, Cook *et al*. 2006b). The importance of these visual cues has led to some successful investigations into reducing *B. aeneus* infestation by changing the colour of the flowers (Giamoustaris & Mithen 1996, Cook *et al*. 2013b) or removing the petals completely (Frearson 2006, Frearson *et al*. 2006). However no preference for a yellow-flowered OSR over an isogenic line with white flowers was found (Frearson 2006, Frearson *et al*. 2006); this may be due to the white-petalled variety used having high reflection in the UV range, as pollen beetles have high sensitivity to UV alongside the green and blue wavelengths that combine to make yellow (Döring *et al*. 2012; Cook *et al*. 2013b). The use of colour cues may explain the findings of Dlabola & Taimr (1965) that beetles can locate OSR fields from at least 300m away, however the second cue – olfaction – may be enough to attract beetles to OSR from that distance (Charpentier 1985, Murlis *et al*. 1992, Moser *et al*. 2009).

OSR produces a suite of volatile compounds (Kjaer 1976, Tollsten & Bergstrom 1988, Robertson *et al*. 1993, Jakobsen *et al*. 1994, Chen & Andreasson 2001, Jönsson *et al*. 2005, Cook *et al*. 2007) and these are known to attract *B. aeneus* (Kirk 1992, Blight *et al*. 1995, Smart *et al*. 1995, Ruther & Thiemann 1997, Smart & Blight 2000, Cook *et al*. 2002, Cook *et al*., 2007, Williams & Cook 2010, Mauchline *et al*. 2013). These attractive compounds are summarised in Williams & Cook (2010) and have been identified as important in host-plant attraction studies using olfactometers (Evans & Allen-Williams 1994, Ruther & Thiemann 1997, Cook *et al*. 2002, Jönsson *et al*. 2005, Mauchline *et al*. 2005), wind tunnels (Cook *et al*. 2006b, 2007) and in the field (Free & Williams

1978, Evans & Allen-Williams 1994, Blight & Smart 1999, Smart & Blight 2000). These compounds have mostly been tested singly however and there is evidence that specific ratios of compounds are used by insects to find their hosts (Bruce *et al*. 2005). These compounds are detected by the beetle via their antennae (Blight *et al*. 1995, Mauchline *et al*. 2008) at least 20m away (Evans & Allen-Williams 1994) and perhaps up to 300m (Taimr *et al*. 1967). Upwind anemotaxis (movement in response to wind stimulus) is used to retain contact with the odour plume from the crop (Evans & Allen-Williams 1994, Williams *et al*. 2007, Williams & Ferguson 2010; Skellern *et al*. 2017). This is supported by the findings of Ferguson *et al*. (2003a, b) that *B. aeneus* mainly colonises crop fields from the downwind edge. In addition, Moser *et al*. (2009) found that the presence of OSR downwind from a study site was a significant predictor of *B. aeneus* density. Visual cues (i.e. presence of yellow flowers) appear to be a strong stimulus for the landing response of incoming beetles (Blight & Smart 1999, Williams *et al*. 2003, Cook *et al*. 2006b, 2007) but the abundance of buds is an important factor in the length of time that the beetles remain on the plant (Frearson *et al*. 2005). *Brassicogethes aeneus* causes most damage to the OSR crop when it is at the green bud stage, and it is at this time, according to Free & Williams (1978) that the beetle numbers are highly concentrated at the edges of the fields, though Mauchline (2003) found no evidence for edge distribution. Ferguson *et al*. (2003a, b) found that *B. aeneus* colonises crop fields from the downwind edge and that the larvae & resultant new generation tended to cluster in the downwind half of a field, the upwind parts remaining relatively uninfested (Williams & Ferguson 2010).

1.2.4 Pollen beetle control and resistance

Since the second half of the 20th Century, control of *B. aeneus* and other pests of OSR in Europe has been predominantly via synthetic insecticides (Thieme *et al*. 2010b). Historically dichlorodiphenyltrichloroethane (DDT) and organophosphates have been used before being phased out (Thieme *et al*. 2010b). Under the current regulatory environment in the EU pyrethroids, indoxacarb, pymetrozine and a limited range of neonicotinoid insecticides are the only synthetic insecticides available, with Chlorpyrifos – an organophosphate approved (until recently) for emergency actions (Thieme *et al*. 2010b, AHDB 2013). All applications of insecticide for *B. aeneus* are as sprays, often in response to populations in the crop exceeding a threshold set by the relevant regulatory body (Thieme *et al*. 2010b), for example the threshold for the UK ranges between 7-25 beetles per plant depending on crop density (HGCA 2013, Ramsden *et al*. 2017), whilst in Denmark it is as low as 1 beetle per plant in the early bud stage (Nielsen & Axelsen 1988b, Hansen 2004) with other European countries generally setting a threshold between 3-8 beetles per plant, depending on growth stage (Richardson 2008; Williams 2010). It is clear that threshold levels for pollen beetles are generally based on little scientific evidence that is often outdated and are frequently ignored by growers (Thieme *et al*. 2010b, Rose *et al*. 2016, Ramsden *et al*. 2017). Given that the cost of insecticides is relatively low this leads to widespread prophylactic use (Thieme *et al*. 2010b, Médiène *et al*. 2011, Ferguson *et al*. 2016, Ramsden *et al*. 2017). For example, the use of pyrethroids as a foliar application has increased since the introduction of fungicides due to the reduction in cost when using tank mixtures with both fungicides & insecticides (Thieme *et al*. 2010b). In Europe, winter OSR typically receives between one and four insecticide applications against *B. aeneus*, with more than 90% of the area sown reported to be sprayed annually (Richardson 2008).

Chemical control brings with it a risk of the target species developing resistance (Edwards *et al*. 2008, Yadouleton *et al*. 2009, Slater *et al*. 2011). Despite insecticidal control in OSR being relatively low in total compared to other crops (Muthmann 2007), OSR does have a consistently high insecticide input that is a central component of the crop protection strategy (AHDB 2013, Hughes *et al*. 2014). The first reports of synthetic insecticide resistance in *B. aeneus* came from Poland in 1965 (Wegorek 1965, IRAC 2006, Thieme *et al*. 2010b), although one report from Germany suggests the possibility of synthetic chemical resistance developing in the late 1910s or early 1920s (Boerner *et al*. 1921). The first incidence of pyrethroid resistance was recorded in northern France in 1999 (Hansen 2003, 2008), with unvalidated anecdotal reports of resistance elsewhere in Europe dating back to 1997 (Thieme *et al*. 2010b). *Brassicogethes aeneus* resistance to a range of insecticides including all classes of pyrethroid is now widespread across Europe (Slater *et al*. 2011, Zimmer & Nauen 2011a, Nauen *et al*. 2012, Ferguson *et al*. 2013, Kaiser *et al*. 2014, Stratanovitch *et al*. 2014, Zimmer *et al*. 2014a, Heimbach & Brandes 2016). Richardson (2008) found a significant correlation at a country level between the number of spray applications of pyrethroids and the year of resistance development in *B. aeneus*. Riggi *et al*. (2016) found a significant effect of the proportion of OSR in the landscape and the mixture of spring and winter OSR on the development of resistance to the pyrethroid lambda-cyhalothrin (Karate®).

There are two main mechanisms of resistance to pyrethroids that have been identified in *B. aeneus* (Nauen *et al*. 2012); one based on modification of the pyrethroid target-site (the sodium channels of the central nervous system (Khambay & Jewess 2005, Nauen 2007)); the second through metabolic detoxification of the pyrethroid from elevated levels of esterases and cytochrome P450 monooxygenases (Vontas *et al*. 2010), i.e. target-site resistance and metabolic resistance, respectively. The target-site mutation is a single amino acid (L1014F) change in the voltage-gated sodium channel (Nauen *et al*. 2012), this confers knock-down resistance (kdr) as the pyrethroid molecules are unable to bind to the sodium channel. Metabolic resistance in *B. aeneus* appears to be through the hydroxylation of the pyrethroid compound by microsomes (Zimmer & Nauen 2011b) mediated by elevated levels of oxidase enzymes containing cytochrome P450 proteins in the CYP2, CYP3, CYP4, CYP6 and mitochondrial subfamilies (Zimmer *et al*. 2014a, b).

Metabolic resistance allows for increased cross-resistance to different pyrethroid compounds than kdr (Zimmer & Nauen 2011b, Zimmer *et al*. 2014a) further reducing control options.

As yet, there is little evidence of neonicotinoid resistance developing in European *B. aeneus* populations (Thieme *et al*. 2010a, Zimmer *et al*. 2014a, Kaiser *et al*. 2015, Milovanović *et al*. 2019, but see Spitzer *et al*. 2020), meaning that neonicotinoid sprays had become an important alternative to pyrethroids in the control of *B. aeneus*. However, due to the potential harmful effects of neonicotinoids on bees, this is unlikely to last and products containing thiacloprid (e.g. Biscaya) will not be permitted for use against pollen beetles from 2021 onwards (but products containing Acetamiprid are still permitted).

The prevalence of resistance in *B. aeneus* to the primary class of insecticide available in Europe, alongside the withdrawal of alternatives, means that steps need to be taken to reduce the risk of economically damaging infestations. An important tactic to reduce the spread of resistant strains is to reduce unnecessary prophylactic spraying of crops (Brown 1981, Brattsten *et al*. 1986). This requires an element of risk on behalf of the grower as not spraying when necessary may result in lower yields, whereas a successful spray when not needed is relatively cheap in costs and labour (Dewar 2017). To encourage growers to reduce prophylactic sprays a range of alternative methods have been, and continue to be, developed. Several methods, such as increased crop rotation (Rusch *et al*. 2011), intercropping or companion planting with pest repellent plants (Endersby & Morgan 1991, Martin *et al*. 2020), trap cropping (Cook *et al*. 2004b, Thöming *et al*. 2020), pest resistant cultivars (Hervé & Cortisero 2016, Seimandi Corda 2019) or reduced tillage (Nilsson 2010) involve a degree of cultural change away from conventional agriculture which may elicit reluctance to adopt from more conservative growers.

One method that requires less cultural change is the use of decision support systems (DSS) that assist the grower in determining whether weather conditions and crop growth stage necessitate crop inspection. The internet has emerged as an important tool for disseminating forecasts and decision support systems (Bajwa *et al.* 2003, Damos 2015); some examples of this are available for *B. aeneus*, including proPlant, the Warwick MORPH model.

The proPlant decision support tool was previously hosted at www.proPlantexpert.com (Johnen & von Richthofen 2013, Ferguson *et al.* 2013), but has been rebranded and available to subscribers only on the Xarvio platform (Hicks 2018). proPlant used accumulated spring temperatures in combination with weather parameters (daily mean & maximum temperatures, wind speed and sunshine hours) to forecast and report on the start and end of *B. aeneus* migration (Johnen *et al*. 2010). Further, the DSS allowed timely assessment of threshold levels of the pest, along with crop growth stage, to determine the need, or not, for control measures. It should be noted this covered a range of pest species beyond *B. aeneus*.

The Warwick University MORPH model, one of a range of models available on the Warwick University Pest Compendium (https://warwick.ac.uk/fac/sci/lifesci/wcc/resources/morphexcel), is a Monte Carlo simulation of the timing of phenological events (spring development, egg maturation, egg and larval development, pupation and pupal development based on thresholds derived from laboratory rearing experiments (Collier *et al.* 2001). The model requires the input of hourly temperature, rainfall and humidity data with the output being a measure of the percentage of emergence of new adults from pupae (Collier *et al*. 2001).

Whilst the proPlant DSS and the Warwick MORPH model for *B. aeneus* could report the timing of important phenological events well (pest migrations and the emergence of adults respectively) (Collier *et al.* 2001, Johnen *et al*. 2010, Ferguson *et al*. 2016), and the Xarvio scouting tool can alert growers of an immediate threat, there are currently no tools that can forecast abundance. For other pest taxa these tools are available, e.g. aphids (Harrington & Woiwod 2007) demonstrating the possibility of developing this aspect of *B. aeneus* DSS tools further.

An aspect of pest forecasting that remains poorly understood in many systems is dispersal, movement and migration (Jeger 1999, Carrière *et al*. 2006, Nathan *et al*. 2008, Huseth *et al.* 2012, Prasad & Prabhakar 2012), both within-field and at a larger scale. Whilst dispersal measures can be inferred from monitoring (Farrell 2000, Osborne *et al*. 2002) or genetics (Ross 2001, Nathan 2001, Jaquiéry *et al.* 2011) and insects are amenable to experimental manipulation to determine dispersal capacity (Chapman *et al.*, 2015), there are still severe challenges in partitioning true dispersal from other ecological factors (Jacobson & Peres-Neto, 2010). It is rare that measures of dispersal distance are used in forecasting pests or decision support tools and when they are present these are often proxies, such as wind speed & direction (Leskinen *et al.* 2011, Burgin *et al.* 2013). Some models that utilise known dispersal parameters do exist, for example Græsbøll *et al.* (2014), who modelled passive and active dispersal in *Culicoides* midges.

1.3 Migration and crop pests

Migration is the seasonal movement of a population of animals from one area to another, usually as a response to changes in temperature, food supply, or the amount of daylight, and often undertaken for the purpose of breeding (Southwood 1962, Reynolds *et al*. 2006, Dingle 2014). For insects this can range from the continent- and generation-spanning migrations of the monarch butterfly (Reppert *et al*. 2016) and the silver Y moth (Chapman *et al*. 2012) to the post-aestivation dispersal of carabid beetles from hedgerows into adjacent cropped habitat in farmland (Fernandez Garcia *et al*. 2000).

1.3.1 Biology of dispersal

In ecology, dispersal is the movement of individuals from one region to another for reproductive purposes leading to gene flow (Wright 1950, Croteau 2010, Dingle 2014). It is generally accepted that there are two primary types of animal dispersal in ecological systems: the movement away from the region of birth to the region where breeding first takes place (natal dispersal) and movement between two successive breeding areas (breeding dispersal) (Clobert *et al*. 2001, Croteau 2010, Dingle 2014). There are other types of movement, for example foraging and migration to non-breeding areas (Dingle 2014), but as these do not directly lead to gene-flow (though they may interact, e.g. Roff & Fairbairn 2001) they are not considered to be dispersal *per se*.

Dispersal is predicated on some form of migratory movement (Kennedy 1985, Dingle 2014). Such movements in insect pests allow the possibility for detection before pest outbreaks become damaging (Pedgley 1993, Day & Knight 1995, Prasad & Prabhakar 2012), or, alternatively drivers of migration can be understood in order to allow forecasting of outbreaks in advance (Day & Knight 1995, Harrington & Woiwod 2007, Prasad & Prabhakar 2012). In *B. aeneus* this migratory dispersal is manifested in their post-winter migration into new breeding areas (natal dispersion) and the movement of females between crop fields (breeding dispersal) (Williams *et al*. 2007, Williams & Cook 2010, Mauchline *et al*. 2013, Juhel *et al*. 2017, 2019).

1.3.2 Modelling approaches

Forecasting migrations requires the modelling of these movements in time and space and requires an understanding of what drives them. Such models can be simple linear regressions of temperature against first flight for example (Harrington & Woiwod 2007), or more complex dynamic models utilising life tables and/or simulations of pest outbreaks under different environmental conditions (Fitt *et al*. 1995, Prasad & Prabhakar 2012). All modelling approaches have advantages and disadvantages. For example, simple linear models may be relatively easy to produce and often outperform more complex models (Rogers 1995, Armstrong 2001, Smith *et al*. 2005) but predicted values are often poor estimates and can struggle with non-linear changes in drivers over time (Smith *et al.* 2005, Oliver & Roy 2015). On the other hand, more complex models may be more accurate and adaptable to changes in drivers over time, but only for a specific location or set of environmental conditions (Smith et al. 2005), or are difficult to parametrise due to the wide range of drivers considered (Oliver & Roy 2015). What unifies all forecasting models is the requirement for data on pest abundance and the drivers of their migrations to base calculations on (Day & Knight 1995, Prasad & Prabhakar 2012).

1.4 Monitoring methods to detect migration

Monitoring pest insects to provide information on migration and abundance is difficult at a landscape scale (Day & Knight 1995, Prasad & Prabhakar 2012). Several methods have proved successful in providing the information growers need to control some pests, but others remain under-researched.

1.4.1. Suction-traps

Suction-traps of the Rothamsted design (Macaulay *et al.* 1988) are 12.2 metres high & draw air through a 244 mm ‘chimney’ at 45 m^3/min. They have operated in the UK since 1964 and a network of these traps now runs throughout the UK, and similar networks of traps of the Rothamsted design or similar operate across Europe (Dedryver *et al.* 2011), North America (Schmidt *et al.* 2012), China (Miao *et al.* 2011) and South Africa (Krüger *et al*. 2014). Originally designed to monitor aphids (Harrington & Woiwod 2007), in recent years their potential to monitor other aerial taxa has become apparent (Shortall *et al*. 2009, Sanders *et al.* 2011).

1.4.1.1 Aphids

The Rothamsted Insect Survey (RIS) celebrated its 50th anniversary in 2014 (Bell *et al.* 2015), for the entire period it has been monitoring aphids and has assembled the most comprehensive standardised dataset on terrestrial invertebrates in the world (Harrington 2014). Alongside their partners in Scotland - SASA (Science and Advice for Scottish Agriculture) – the RIS run a network of traps in England and Scotland. In addition to providing grist for a range of seminal papers in a range of disciplines (Harrington 2014, Bell *et al*. 2015) the RIS also provides data to growers on the timing and size of aphid migrations (Harrington *et al.* 2012). Currently these data are provided as a series of weekly bulletins that provide timely information for farmers, growers and agronomists throughout the critical periods in the growing season, with commentary provided to elucidate the raw numbers.

1.4.1.2 Other taxa

All samples taken by RIS are stored; as yet there are no other groups or species currently monitored quantitively on a regular basis from the suction-traps, though weekly reports on economically important insects noted in the traps are available to subscribers. There is potential to monitor species of economic importance, including *Culicoides* biting midges (Diptera: Ceratopogonidae), vectors of bluetongue disease that affect livestock (Fassotte *et al.* 2008, Sanders *et al.* 2011), and two serious pests of brassica crops *Psylliodes chrysocephala* (Coleoptera: Chrysomelidae) and *B. aeneus* (this thesis).

1.4.2 Radar entomology

The detection of insect migrations using various radar technologies such as vertical-looking radar (VLR) (Chapman *et al*. 2003); scanning entomological radar (SER) (Riley *et al*. 1995), tracking entomological radar (TER) (Westbrook *et al*. 1995), airborne entomological radar (AER) (Hobbs and Wolf 1989); or weather radar (Bauer *et al*. 2017) is an emerging and increasingly useful technology (Chapman *et al*. 2003, Leskinen *et al*. 2011, Bauer *et al*. 2017, Abd El-Ghany *et al*. 2020, Kelly & Stepanian 2020). VLR, TER and SER can be deployed at specific sites to provide an assessment of the migration of larger insect pests across a wide area (Chapman *et al*. 2003, Reynolds *et al*. 2006, Bauer *et al*. 2017, Pearson *et al*. 2020), a key strength of VLR being the ability to monitor at different altitudes simultaneously to give a profile of insect migration allowing for short-term forecasts (Chapman *et al*. 2003, Abd El-Ghany *et al*. 2020). In a similar manner, weather radar can be used to track insect movement (Chilson *et al*. 2012, Stepanian *et al*. 2016, Kelly & Stepanian 2020) and be used to form the basis of a pest migration alarm system (Nieminen *et al*. 2000, Leskinen *et al*. 2011, Boulanger *et al*. 2017). However, it is difficult to monitor smaller sized insects using these radar technologies.

1.4.3 In-field sampling

There are several methods that are available to monitor pest migration in-field (Southwood & Henderson 2009, McCravy 2018), ranging from direct inspection of plants to trapping pest species to more recent technologies for remote detection of pests.

1.4.3.1 Plant scouting

The simplest method for monitoring pests in fields involves visually inspecting (“scouting”) the crop for infestation (Pedigo & Buntin 1993). Various methods of inspection can be employed, for example, counting insects or their eggs on plants (Mitchell *et al*. 2000); beating individual plants into trays to count the number of insects present (Tuovinen & Parikka 1996, Walters *et al*. 2003); sweep-netting to gain an estimate of insect abundance (Morandin *et al*. 2014, Cato *et al*. 2019); determining the percentage of plants infested (Cartwright *et al*. 1987); or assessing the amount of damage to plants by pests (Mitchell *et al*. 2000). *Brassicogethes aeneus* is generally assessed by the beating tray method (Walters *et al*. 2003, AHDB 2013), with transects walked into the crop and random plants selected for beating at set intervals (AHDB 2013). Fera Science Ltd. (Fera), formerly the Food and Environment Research Agency, monitors a range of crops across the UK utilising this technique, including OSR for *B. aeneus* and *Psylliodes chrysocephala* – transects are taken through target fields and crop samples examined for presence of adults (*B. aeneus*) or larvae (*P. chrysocephala*) and damage. Data are disseminated via reports and through the trade press (AHDB 2013) and online via https://secure.fera.defra.gov.uk/cropmonitor/wosr/surveys/wosrPestAssLab.cfm , the CropMonitor website (Collins *pers. comm*).

1.4.3.2 Water traps

Water traps (Moericke 1951) are frequently deployed to monitor in-field pests. Yellow water traps in particular are known to preferentially attract a range of crop pests, including aphids (Northing & Dale 2009) and pollen beetles (Williams 2010). Fera operate water-traps around the country, primarily in potato fields to monitor aphids. Data are disseminated to registered users in the form of alerts on the first *Myzus persicae* record in the season and once potato virus index values reach a certain threshold (Northing & Dale 2009).

Alongside their contribution to the UK suction-trap network, SASA also carry out yellow water tray trapping in potato crops throughout Scotland with aphids the main target (Highet *pers. comm.*). Results from the suction-trap are summarised in a weekly bulletin with interpretation available to subscribers. Water-trap results are passed on to growers in the form of weekly updates on the SASA website, with additional work on virus transmission risk also provided (Highet *pers. comm.*).

Yellow water tray traps in sugar beet are run by the British Beet Research Organisation in the major sugar beet growing areas – predominantly in Eastern England – to monitor aphid populations. Results are available as weekly bulletins, with summaries in the trade press (https://bbro.co.uk/publications/beet-review/). In addition, the AHDB report online on horticultural pests in water tray traps based at the Warwick Crop Centre station at Wellesbourne (https://blogs.warwick.ac.uk/rosemarycollier/).

As part of their DSS package (Hicks 2018), Xarvio offer automated yellow trap analysis as part of their scouting package (https://www.xarvio.com/gb/en/products/scouting/yellow-trap-analysis.html). Growers are able to identify and count harmful insects in their traps using the app and can share their results with other users to gain a general impression of the insect pressure in their region.

1.4.3.3 Sticky traps

Sticky traps (Broadbent 1948) consist of a sheet, cylinder, or globe covered in glue and mounted at varying heights, either in or around the crop. Sticky traps have been used to monitor a range of pests, including fruit flies (Cytrynowicz *et al*. 1982), whiteflies – both in greenhouses (Webb & Smith 1980) and in the field (Atakan & Canhilal 2004) – cornworms (Midgarden *et al*. 1993) and leafhoppers (Chancellor *et al*. 1997).

In the UK, the use of sticky trap monitoring for use within warning systems in agricultural crops is limited to carrot flies *Chamaepsila rosae* (Fab.) (Collier *et al*. 1990), with monitoring results and weather-based forecasts distributed via the AHDB Pest Bulletin (https://ahdb.org.uk/knowledge-library/ahdb-pest-bulletin). *Brassicogethes aeneus* and other pests are recorded and reported as bycatch but are not the target pests (https://blogs.warwick.ac.uk/rosemarycollier/). In addition, sticky traps are used as a research tool for understanding the phenology and distribution of pest and natural enemy species (e.g. Heathcote 1957, Finch & Collier 1989, Collier *et al*. 1990, , Longley *et al*. 1997, Bruce *et al*. 2005), including *B. aeneus* (Skellern *et al*. 2018).

1.4.3.4 Pheromone traps

Despite being known for many decades prior (e.g. Fabre 1879, Lintner 1882), and posited even earlier (Butler 1609, Darwin 1896, Wyatt 2009), pheromones were named as such in 1959

(Karlsen & Lüscher 1959) after the first chemical identification was made by Butenandt (1959). Pheromones are a suite of chemicals that are excreted by individuals that elicit a behavioural or endocrine response from individuals of the same species (Hardie & Minks 1999, Doty 2010), although they may also influence the behaviour of other species such as allowing host location by parasitoids (Feener Jr. *et al*. 1996, Du *et al*. 1998). Utilising pheromone attractants to trap insects was first posited in the early 1960s (Coppel *et al*. 1960, Jacobsen & Beroza 1964), with a range of traps developed for different target insects (Elkington & Cardé 1984, Wyatt 2009). These traps utilise the attractant potential of (usually) female sex pheromones to attract the target species and either include sheets covered with glue or, more rarely, a non-lethal means of collection (Elkington & Cardé 1984) in order to monitor the target species.

In the UK, pheromone traps have been used to monitor a range of insect pests, including the orange wheat blossom midge (*Sitodiplosis mosellana* (Gehin)) (Bruce *et al*. 2007), the saddle gall midge (*Haplodiplosis marginata* (von Roser)) (Rowley *et al*. 2017), the oak processionary moth (*Thaumetopoea processionea* L.) (Williams *et al*. 2013), orchard moths (*Cydia pomonella* (L.) and *Archips podana* (Scopoli)) (Alford *et al*. 1979). As yet, only a potential epideictic (repellent) pheromone has been identified in *B. aeneus* (Ruther & Thiemann 1997, Cook *et al*. 2006a), however there has been progress in developing a sticky-trap baited with the plant volatile phenylacetaldehyde, known to attract *B. aeneus* (Cook *et al*. 2013a). This trap has been deployed around the UK allowing growers to make their own assessments about *B. aeneus* infestation rates within their crops.

1.4.3.5 Other technologies

One aspect that unites both in-field sampling and observation through trapping methods is the large amount of time required for counting and identifying samples. This includes the time spent collecting and quantifying samples and the training required in order to carry out these tasks (Southwood & Henderson 2009, Ahmad *et al*. 2018, Preti *et al*. 2020). There are several automated techniques that can detect species of interest without the input of a human observer. These techniques include automated traps with infra-red (Hendricks 1985, Ahmad *et al*. 2018), optical (Hendricks 1990, Doitsidis *et al*. 2017, Pearson *et al*. 2020, Preti *et al*. 2020), acoustic (Potamitis *et al*. 2017, Sandrini Moraes *et al*. 2019) or lidar (Mei *et al*. 2012, Ahmad *et al*. 2018, Jansson & Brydegaard 2018, Kirkeby *et al*. 2021) sensors to relay pest occurrence information to growers, either directly or via a data processing centre. As yet there are few applications of these novel technologies relating to *B. aeneus* and thus further discussion of these techniques is outside the remit of this thesis.

1.5 Pest forecasting

The monitoring of pests has a major benefit for growers: the ability to forecast future outbreaks, spatially and/or temporally (Bardner *et al*. 1982). There are many approaches to best determine the location, timing and extent of pest outbreaks in crops. Many utilise weather factors, for example air and soil temperature (Junk *et al*. 2016, Johnen *et al*. 2010), rainfall (Thackray *et al*. 2009, Veran *et al.* 2015), mean temperature in autumn (Thackray *et al*. 2009) or winter (Yamamura *et al*. 2006, Eickermann *et al*. 2014, Matsumura *et al*. 2015), maximum temperature of the previous year (Hameed *et al*. 2015), day degrees (Wilson & Barnett 1983, Tu *et al*. 2014), sunshine hours (Junk *et al*. 2016) or relative humidity (Hameed *et al*. 2015). Others consider land cover (Veran *et al*. 2015) and a few combine a range of weather variables with time of year to account for pest development (Sanders *et al*. 2011, Hirschi *et al*. 2012). Most forecasting measures also rely on accurate measurements of the pest abundance in previous years (Yamamura *et al*. 2006, Hameed *et al*. 2015, Matsumura *et al*. 2015, Veran *et al*. 2015).

The lack of formal pest information forecasts and related decision support systems across production systems obstructs the development of improved pest management (McBratney *et al*. 2005, Damos 2015). A further complication is the changing agricultural landscape, technological advances and climate change creating a shifting context for solving pest management problems (Strand 2000).

1.6 Aims and objectives of this study

This study aims to fill some of the knowledge gaps in our understanding of *B. aeneus* population dynamics. Namely: their long-term population dynamics; the drivers of annual and daily abundance as estimated by trap catches; the spatial distribution and their dispersal ability and population structure in the UK. This knowledge is needed to develop models using historic and current suction-trapping data to form predictions of pollen beetle migration into both winter and spring oilseed rape crops over a useful time frame for farmers. These models should also provide estimates of the size of the migrating population. Together these outputs can be built into a decision support system (DSS) for farmers so that they can more accurately assess and manage the risk of pollen beetle damage to their oilseed rape crops and therefore reduce unnecessary insecticide applications.

The aims were addressed via a series of objectives:

1) Use suction-traps to define the environmental and metrological drivers of migration in *B. aeneus* to enable prediction of timing of migration (Chapters 2 & 3).

2) Investigate population dynamics of *B. aeneus* to predict the size of the immigrating population (Chapters 2 & 3).

3) Investigate the relationship between suction-trap catch and population abundance in crops to determine whether suction-trap catches can inform field observations or provide insight into *B. aeneus* spatial distribution in England (Chapter 4).

4) Use flight-mills and molecular data to investigate the dispersal ability and population structure of the English *B. aeneus* population(s) (Chapter 5).

Chapter 2.

Long-term trends in *Brassicogethes aeneus* abundance and oilseed rape cultivation.

Chapter Summary

Data on pollen beetle (Brassicogethes aeneus) *numbers caught in Rothamsted suction-traps from 1987-2018 were investigated to determine long-term trends in annual counts and their relationship with area of oilseed rape* (Brassica napus) *grown at a national and regional level. The trend at seasonal and monthly levels were compared to the overall trend, and the relationship between counts in late summer of newly-emerged beetles and counts in spring/early summer of overwintered beetles the following year and the inter-generational relationship within year were also investigated.*

It was found that, although both B. aeneus *counts and area of oilseed rape grown in the UK increased over time, annual counts of B. aeneus were negatively related to the area of oilseed rape grown nationally that year. Conversely, the relationship was positive when oilseed rape area was considered on the regional level. The increasing trend in* B. aeneus *annual counts was also shown in most seasons and months except for March/April and July which showed no overall trend across the time period studied. In addition, it was found that counts of beetles from June 1st onwards (the newly emerged generation) showed a positive relationship with counts up to June 1st the following year, indicating that it may be possible to forecast the abundance of the spring migration of* B. aeneus *a few months in advance utilising suction-trap samples.*

2.1 Introduction

2.1.1 Long-term Monitoring

2.1.1.1 What is long-term monitoring?

Long-term monitoring is the study of a system over an extended period of time using repeatable, standardised sampling methods (Callahan 1984, Franklin 1989, Lindenmayer & Likens 2010, Magurran *et al*. 2010, Johnson 2012, Lindenmayer *et al*. 2012). The definition of this extended period varies between systems and the hypothesis being tested; for insects the temporal scale and extent can vary from a few months (Englund 2001) to several decades (Gibbs 2000, Szentkirályi 2002, Hughes *et al*. 2017, Bell *et al*. 2020). Long-term monitoring in ecological systems allows researchers to assess trends in abundance, phenology and/or distribution of the target species or habitats over time (Storkey *et al*. 2016, Hughes *et al*. 2017) and allows for the development and testing of ecological theories (Hanski & Woiwod 1993, Lovett *et al*. 2007, Gitzen *et al.* 2012). This in turn allows the study of population dynamics (Franklin 1989, Gibbs 2000, Doak *et al*. 2008) and gives a greater knowledge of the effects of environmental change (Willis *et al*. 2007, Clutton-Brock & Sheldon 2010, Magurran *et al*. 2010, Hughes *et al*. 2017) and interventions (Bell *et al*. 2012) on populations and habitats. The knowledge gained can then allow for remedial action in the case of species of conservation concern (Legg & Nagy 2006, Lovett *et al*. 2007, Willis *et al*. 2007, Magurran *et al*. 2010, Carpenter *et al*. 2011, Lindenmayer *et al*. 2012), or inform control options of economic pest species (Bell *et al*. 2012, Wang *et al*. 2019).

2.1.1.2 Long-term insect monitoring and forecasting in agriculture

Monitoring agricultural pest systems allows us to forecast future outbreaks over space and time (Bardner *et al*. 1982, Finch *et al*. 1996). Currently, most insecticide application decisions are made either prophylactically or based upon economic thresholds of pest presence on inspection (Teng 1994, Lima *et al*. 2014, Onstad *et al.* 2019), however these are not economically viable in all cropping systems (Reisig *et al*. 2012, Ramsden *et al.* 2017) and, indeed, thresholds are often inaccurate and not widely used (Ramsden *et al*. 2017). Predictions generated from forecasts can be a major benefit if this knowledge is disseminated to growers, allowing them to take preventative rather than prophylactic measures to control the target pest (Teng 1994, de Freitas Bueno *et al*. 2011, Holzworth *et al*. 2015), provided that they are issued within a generous treatment window. These types of forecasts are often referred to as decision support systems (DSS) when the prediction is linked to a possible intervention to exercise control over a pest (Coulson & Saunders 1987, Lagos-Ortiz *et al*. 2018, González-Andújar 2020). DSSs are a crucial tool in Integrated Pest Management (IPM) cultivation (Teng 1994, Kogan 1998, Lima *et al.* 2014, González-Andújar 2020), which has been shown to reduce pesticide use (Teng 1994, de Freitas Bueno *et al*. 2011) and improve economic performance (Rola & Pingali 1993) in certain cropping systems.

2.1.1.3 The Rothamsted Insect Survey

There are very few datasets which allow long-term quantitative trends in insect populations to be derived (Montgomery *et al*. 2019, Saunders, 2019). Perhaps the best is the Rothamsted Insect Survey (Bell *et al*. 2015) database and archives. The database contains >50 M records of, primarily, moth and aphid abundance data over more than 50 years across the UK from a network of light traps and 12.2m suction-traps (Harrington 2014, Storkey *et al*. 2016), whilst the archive contains

~200,000 samples of bycatch from 39 suction-traps (Shortall 2013). This long-term dataset has enabled forecasting tools to be developed for some species of aphid that show a strong relationship between winter temperature and their phenology and abundance (Harrington & Woiwod 2007, Bell *et al*. 2015) and has also shown the potential for forecasting in other taxa found in the bycatch, for example biting midges (White *et al*. 2017). The bycatch has recently been demonstrated to have the potential to inform studies of a range of aerial taxa (Shortall *et al*. 2009, Sanders *et al*. 2011, Pérez-Rodríguez *et al.* 2015), however, previous to this work, pollen beetles (*Brassicogethes aeneus* (Fab.)) had not been studied in detail using the 12.2 m suction-trap network.

2.1.2 Pollen beetles

2.1.2.1 Pollen beetles and oilseed rape

Pollen beetles (*Brassicogethes aeneus* (Fab.)) are a pernicious pest of oilseed rape (OSR) and other brassicas (Williams & Free 1978, Finch *et al*., 1996, Williams 2010), causing feeding and oviposition damage in buds causing them to abscise (Free & Williams 1978, Ekbom & Borg 1996, Williams 2010). As most winter OSR is past the susceptible stage by the time the beetles emerge from hibernation in the spring the damage to these crops is usually minimal (Williams 2010), although early emergence or late flowering can lead to economic loss (Zlof 2008). In spring OSR and horticultural crops, however, damage can be extensive (Finch *et al*. 1990, 1996, Alford *et al*. 2003, Hansen 2004). Thus, it is of interest to both researchers and growers to measure seasonal and overall abundance of *B. aeneus* across years. This may then give insight into how the populations fluctuate through time with changes in climate and land-use and allow the development of forecasting and DSS tools.

2.1.2.2 Pollen beetle habitat has increased through time

Much of the recent research on long-term insect trends has focussed on declines in abundance, with habitat loss often cited as an important driver of these declines (e.g. Biesmeijer *et al*. 2006, Conrad *et al*. 2006, Van Dyck *et al*. 2009, Dirzo *et al*. 2014, Fox 2013, Spiller & Dettmers 2019, Saunders 2019). Conversely, in the UK and northern and eastern Europe, *B. aeneus* has seen its area of habitat increase in recent years (Hokkanen 2000, Gehringer *et al*. 2007). The development of reduced glucosinolate and eruic acid varieties of OSR in Canada in the 1970s (Bunting 1984, Booth & Gunstone 2004) led to a rapid rise in the area under oilseed cultivation in the UK (Fig. 2.1). The crop briefly became the second most widely grown crop in the UK, reaching a peak in 2012 at 756,000 hectares with more than 600,000 hectares grown in most years – representing around 9-12% of the total cropped area in the UK (DEFRA, 2014, 2018). This increase in acreage was reflected across most of northern and eastern Europe (Gehringer *et al*. 2007). It has already been demonstrated that *B. aeneus* is able to rapidly shift to OSR from native brassica hosts when these crops are introduced to an area (Hokkanen 2000). This ability to shift to OSR allowed the beetle to exploit the increase in acreage that occurred in Europe in the second half of the 20th century and become an abundant pest throughout the region (Hokkanen 2000). There are indications that area increase has plateaued since 2012, UK-wide sown area has declined in all years except 2018, along with yield in most years (DEFRA 2018, 2019a). The crop is now at its lowest acreage in a decade (AHDB).

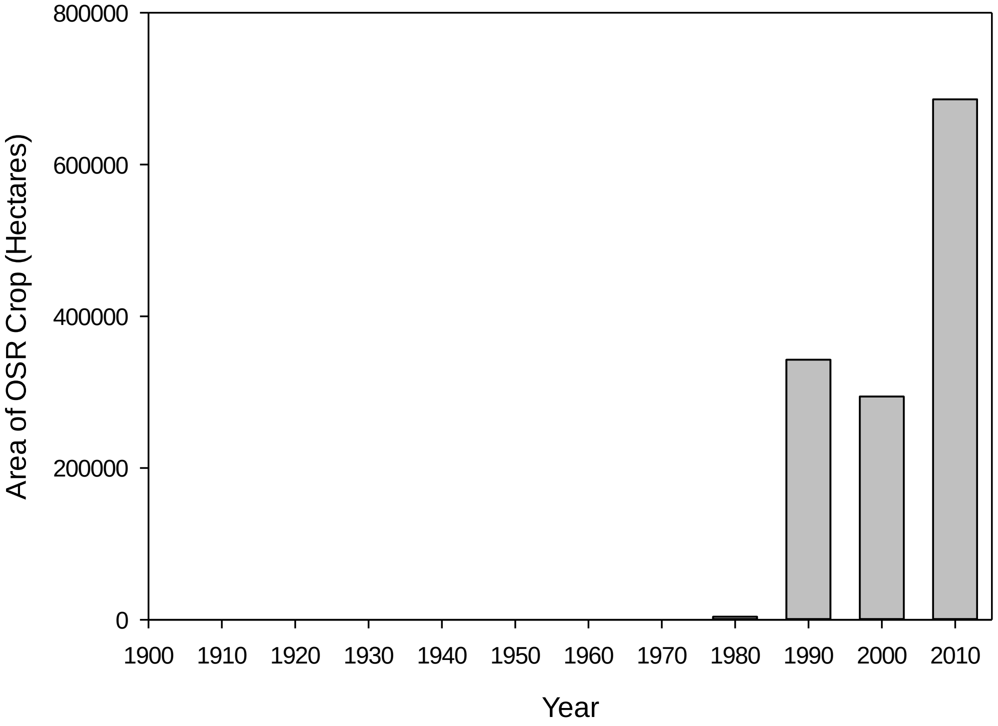


Fig. 2.1: **DEFRA June Census statistics for oilseed rape in the UK 1900-2010. The crop was not grown in sufficient quantity to appear in the statistics until 1970 and has since risen to become an important UK crop.**

2.2 Aims and objectives

The main aims of this study were to understand the long-term pollen beetle population trends in the UK and to elucidate the meteorological and landscape level drivers of these trends in order to assess the potential for forecasting beetle abundance using suction-traps.

The objectives were:

1) to use suction-trap catches to visualise annual and daily flight patterns of *B. aeneus* over more than thirty years using data from two suction-traps in the south-east of England and over four years using data from multiple suction-traps across England.

2) to use statistical methods to compare the trends in *B. aeneus* population dynamics to changes in the area of OSR cultivation.

3) to determine any meteorological drivers of change in the *B. aeneus* flight patterns.

2.3 Materials and Methods

2.3.1 Suction-traps

Rothamsted Insect Survey (RIS) suction-traps (Macaulay *et al*. 1988, Bell *et al*. 2015) have been used to monitor and forecast aphids in the UK since 1964 (Harrington & Woiwod 2007). The traps sample aerial insect populations at 12.2m above ground level at 45 m^3/min (Macaulay *et al*. 1988); samples are collected daily with the aphids removed, identified and counted and the bycatch stored in a mixture of ethanol & glycerol (Shortall *et al*. 2009). The presence of glycerol means that, though samples do become dehydrated over time, they remain undamaged on rehydration.

2.3.1.1 Beetle Counts

For the duration of this study the RIS suction-trap network in England consisted of 12 sites (Fig. 2.2). Daily or weekly samples from all suction-trap sites between January and September (the

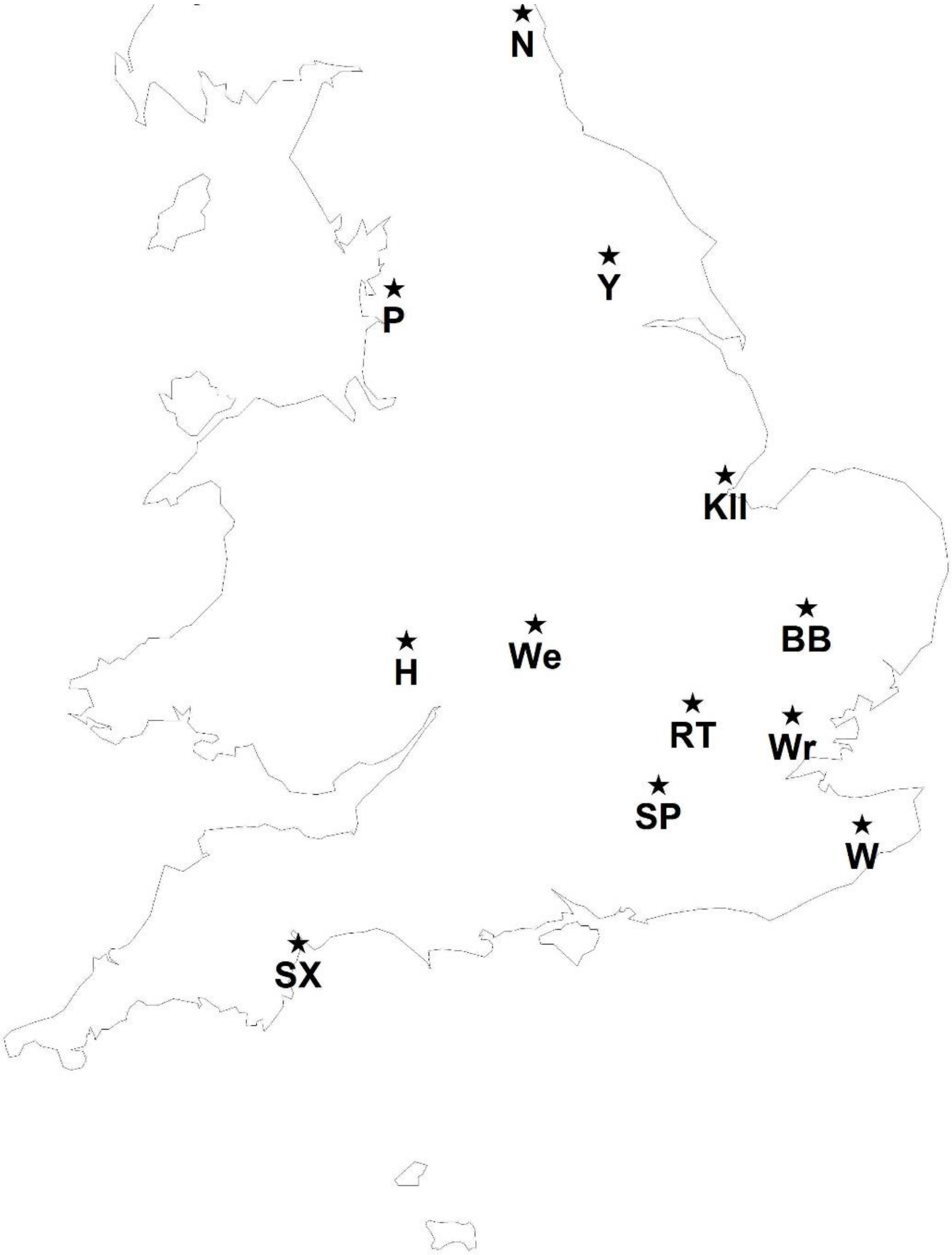


Fig. 2.2: The Rothamsted Insect Survey suction-trap network in England. Site codes, from north to south are: N: Newcastle, Y: York, P: Preston, KII: Kirton II, BB: Broom's Barn, We: Wellesbourne, H: Hereford, RT: Rothamsted Tower, Wr: Writtle, SP: Silwood Park, W: Wye, SX: Starcross.

known flight period of *B. aeneus* in the RIS samples) for the years 2015-2018 were sorted and *B. aeneus* were identified and counted, with the exception of Preston, where, due to low beetle counts in high volume samples, only 2015 samples were investigated. In addition, two sites were selected for investigation of the long-term trends of *B. aeneus* abundance based on trap longevity, consistency of operation and known high counts: Rothamsted Tower 1987-2014 and Broom's Barn 1989-2014; *B. aeneus* from these site years were counted in the same way as the 2015-18 samples. Due to the time needed to process samples, no more than two sites could be used for the long-term part of this study. *Brassicogethes aeneus* is difficult to separate quickly from the closely related *B. viridescens* (Fab.), however the latter is far less common (Billqvist & Ekbom 2001, Metspalu *et al*. 2011, Ouvrard *et al*. 2016) thus, all pollen beetles were assumed to be the former.

To assess the interaction between generations within the year and the effect of overwintering on beetle numbers the annual total of beetles caught at the trap was subdivided for each year at June 1st. Those caught between the start of the year and June 1st were assumed to be those emerging from hibernation and migrating to breeding sites (the Overwintering Beetles), with those caught from June 1st onwards assumed to be those of the new generation and labelled Newly Emerged Beetles.

2.3.2 Oilseed rape data sources

It is difficult, if not impossible, to determine the precise locations and area of oilseed rape cultivation in the vicinity of the suction-trap sites, particularly historically. Instead, regional-level data were sourced. Data for area of OSR sown (ha), yield (t/ha) and production (t) from 1983-2019 (total OSR) and 1983-2015 (winter OSR) for the UK as a whole and 2001-2015 for English regions were accessed from the Defra statistics portal (https://www.gov.uk/government/statistical-datasets/structure-of-the-agricultural-industry-in-england-and-the-uk-at-june date last downloaded 10/01/2010) and data.gov.uk (https://data.gov.uk/dataset/76ca636f-a449-44ba-ac2f-f8febecec2a2/cereals-and-oilseeds-production-harvest date last downloaded: 09/01/2020), respectively.

2.4 Analyses

2.4.1 Statistics

All analyses were carried out in R versions 3.5.0 and 3.6.1 (R Core Team 2018) using RStudio Version 1.1.453 (RStudio Team 2015). *A priori* the statistical distribution of all observations (daily count of *B. aeneus* and annual values for area of OSR) was assessed using the *fitdistrplus* library along with the *post-hoc* model diagnostics in the routine gam.check, the negative binomial or the quasipoisson distributions proved to be the best fit to the data (Delignette-Muller & Dutang 2015, Wood 2017).

The R package *poptrends* (Knape 2016) was used for trend analysis of the data, *poptrends* is an extension of the *mgcv* library (Wood 2017) for generalized additive models (GAMs) and generalized additive mixed models (GAMMs).

The following analyses were carried out:

i) an analysis of the long-term trends of OSR cultivation in the UK as a whole and the south-east region.

ii) an assessment of the long-term trends in annual *B. aeneus* counts at Rothamsted and Broom’s Barn.

iii) an assessment of the relationship between annual *B. aeneus* counts and the area sown of OSR both at a national scale and at a scale local to the traps.

iv) an exploration of how *B. aeneus* counts change between years and generations.

2.4.1.1 Generalized Additive Models

In ecology, it is often the case that the relationship between explanatory variables and observations are non-linear. In these circumstances, rather than using linear regression, it is preferable to utilise smoothing models such as generalized additive models (GAMs) that allow for non-linear relationships between observations and multiple explanatory variables (Zuur *et al*. 2009).

2.4.1.1.1 Smooth functions

GAMs can be defined as a generalized linear model with a smooth function. These smooth functions, splines (piecewise polynomial functions) in the *mgcv* package used here, can be linear but are more usually allowed to be flexible (wiggle) to give a better fit to the data. A GAM allows a spline to be fitted to the data by, effectively, splitting the fitted curve into separate sections of polynomial smoothers delineated by knots (Wood 2017). Each section of smoothed curve has a different coefficient, but they match with their neighbours at the knot points (Wood 2017).

2.4.1.1.2 Knots

Knot points are essentially a mathematical function to ensure that the smoothers are neatly connected (Zuur *et al*. 2009), they can be estimated from the data but it is argued (Zuur *et al*. 2009, Wood 2017) that they should be fixed at a sufficiently large number to capture the maximum wiggliness required and use penalised splines and cross-validation to estimate the shape of the relationship. Here, knots occur at the annual or monthly period along the trend.

2.4.1.1.3 Penalised Splines

Penalised splines reduce overfitting by introducing a penalty to the amount of wiggle in the data and forcing the model to minimise this penalty whilst still optimising the fit. This optimisation allows the model to find the best compromise between wiggliness and smoothness for each section of the data and thus reduces overfitting.

2.4.1.1.4 Cross-validation

Cross-validation refers to a suite of mathematical techniques that drop observations from a dataset, estimate the smooth terms from the remaining *n-1* observations, calculate the value of the dropped observation, compare the observed value to the calculated value, then repeat for all observations in the dataset giving an estimation of the prediction error of the model (Zuur *et al*. 2009, Wood 2017). Estimation of the smoothing parameter can then be based on the mean square prediction error from the cross-validation process with the smoothed fit being the one that minimises the prediction error (Wood 2017).

2.4.1.1.5 Restricted Maximum Likelihood

The suggested cross-validation method to use for GAMs is Restricted Maximum Likelihood (REML) (Wood 2011, Simpson 2018). This approach measures the fit of the smoothing parameters across the (scaled) average of the likelihood over all values of the variable in question (Wood 2017). Since this approach does not use any degrees of freedom in the model by estimated fixed effects (as found in other cross-validation methods), it tends to avoid under-estimating the level of smooth required, allowing for a more flexible model than other methods (Wood 2017).

2.4.1.2 Generalized Additive Mixed Models

A restriction of GAMs is that they assume that data points are independent, an assumption violated by time-series data. One way of accounting for correlation between observations is to use mixed models. As GAMs are an extension of GLMs with a smooth parameter, so are generalized additive mixed models (GAMMs) an extension of generalized linear mixed models (GLMMs) with smoothing (Lin & Zang 1999). As with GAMs (described in section 2.4.1.1 above), GAMMs use nonparametric regression to allow the explanatory variables to have a flexible functional relationship with observations while accounting for correlation between data points by using random effects (Lin & Zang 1999). In all models used here, site was included as a random effect component which allowed a region-wide spatial scale to be adopted.

2.4.1.2.1 Mixed Models

A mixed model contains both fixed effects and random effects in assessing the relationship between observations and explanatory variables. In a linear regression without autocorrelation the explanatory variables are fixed effects, while the observation data can be assumed to be random because the data are drawn from a larger population and, in essence, represent a random sample from that population. The resulting variability between data points (or residual variance) allows the calculation of the amount of unexplained error in the regression (for example residual mean squared error). In autocorrelated data, such as time-series data, in addition to the random variability of the observations there may also be random variability in the explanatory variables. Mixed models account for this aspect of variability by using random effects.

2.4.1.2.2 Random Effects

Random effects come in two forms, which can be combined or computed separately: simple random intercepts allow observations to be higher or lower for each level of explanatory variable and random slopes allow the strength of each fixed effect to vary for each level of explanatory variable. This statistical infrastructure allows the model to derive inference about the larger population and incorporate the variability in explanatory variables to better describe how they relate to observations as fixed effects.

2.4.1.3 Model Selection

Given the high number of variables under consideration at times it was necessary to determine which model of a range of choices was the best to use. I used the Akaike Information Criterion (AIC) (Akaike 1973, Sakamoto *et al.* 1986) method for model selection in these cases. AIC is a now-standard technique that gives the best compromise between the goodness of fit of the model and its simplicity by estimating the amount of information lost by a given model relative to other, comparable, models using the same data. The less information lost by a model, the better that model is. AIC was estimated using the AIC() function in base R.

2.4.2 Oilseed rape trends over time

To ascertain that the data for regional and national trends in OSR production over time matched those reported, and were amenable to analysis, a smoothed GAMM model was fitted with a quasipoisson distribution and automatic selection of optimal degrees of freedom of the cubic smoothing spline (see Appendix A: Chapter 2).

2.4.3 Long-term trends in pollen beetle abundance at Rothamsted and Broom's Barn

Annual *B. aeneus* population trends for the East of England, using Rothamsted and Broom's Barn data combined, alongside some seasonal subsets of the data as defined in Table 2.1, were estimated using poptrend. A spline was fitted with a negative binomial distribution for overdispersed data and automatic selection of optimal degrees of freedom. The dispersion parameter theta (Θ) was estimated at 3.707052 after estimation using the *nb* function of mgcv. See Appendix A, Chapter 2 for the R code for this step. Year and site were set as a random effect and the trend was smooth, rather than loglinear or index and the distribution was a negative binomial.

The poptrend package has an explicit test for short-term trend estimates along the overall trend. If significant, these short-term trends are depicted in orange for decline and green for an increase shown on top of the long-term trend coloured black (Figs. 2.8, 2.11-2.13, 2.15-2.18). Further information can be found at https://github.com/jknape/poptrend.

Season/Month	Duration	theta (Θ)
Spring	1st March – 31st May	3.961196
Summer	1st June – 30th September	4.163855
March/April	1st March – 30th April	3.305895
May	1st – 31st May	4.361432
June	1st – 30th June	3.521173
July	1st – 31st July	2.858764
August/September	1st August – 30th September	0.695337

Table 2.1: Subsets of *Brassicogethes aeneus* count data analysed in poptrends alongside the full dataset with calculated theta (Θ) values. March/April and August/September were combined due to generally low counts in those months.

2.4.4 Temporal autocorrelation

Temporal autocorrelation is the correlation of population phenomena in time, in this case the behaviour of the timing of *B. aeneus* abundance was analysed using mgcv and itsadug (van Rij *et al*. 2017), a GAM was run as detailed below and the autocorrelation function (ACF) of the residuals was checked using the function *acf_resid*. Whilst there was an indication of a seasonal component in the residuals at both sites, notably with a lag of two years (Fig. 2.3), however this seasonal component was dropped by the GAM procedure, likely as a result of overfitting.

2.4.5 Relationship between number of pollen beetles and area of OSR at different temporal and spatial scales

The relationship between *B. aeneus* counts and area of OSR sown in the UK, at local and national scale was assessed using a GAM within the mgcv package (Wood 2017). See Appendix A Chapter 2 for detail on how the data were modelled. A cubic spline was used for the Area and Year terms, the distribution was negative binomial with a log link specified. The smoothing method was set as Restricted Maximum Likelihood (REML), which tends to avoid under-smoothing (Wood 2017).

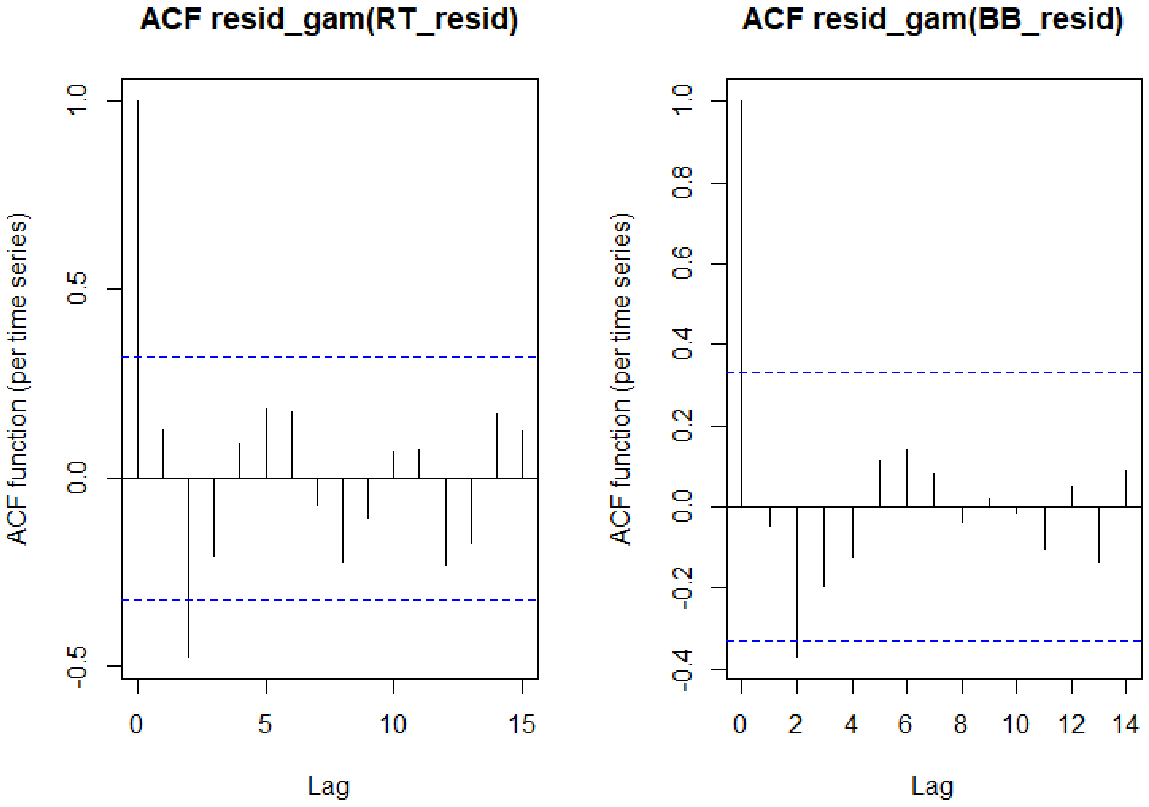


Fig. 2.3: Autocorrelation Function (ACF) plots of *Brassicogethes aeneus* daily count residuals for Rothamsted Tower (RT) (left) and Broom's Barn (BB) (right). Both sites show some sign of a lagged correlation at year 2 (further investigation revealed this to be unimportant in the models).

For the shorter time period (2001-2015) where regional data were available for OSR area, Pearson correlation analyses were performed using the *cor* function in base R and plotted using the *ggscatter* function from the ggpubr library. Data from all sites from 2015 were compared with winter OSR area sown for the region (Fig. 2.4, Table 2.2). Correlation analysis was preferred over more complex models such as GAMs here due to the reduced power from the restricted (one year) dataset. Pearson correlation analyses were performed using the *cor* function in base R and plotted using the *ggscatter* function from the ggpubr library (Kassambara 2019).

Site	**latitude**	**longitude**	**Region**
Newcastle	55.213254	-1.685083	North East
York	54.014616	-0.973205	Yorkshire
Preston	53.854383	-2.766990	North West
Hereford	52.124201	-2.638156	West Midlands
Wellesbourne	52.205975	-1.605017	West Midlands
Kirton	52.924454	-0.052153	East Midlands
Broom's Barn	52.260681	0.568430	Eastern Area
Rothamsted	51.806997	-0.360091	Eastern Area
Writtle	51.733599	0.429233	Eastern Area
Silwood Park	51.409410	-0.643357	South East
Wye	51.185507	0.944941	South East
Starcross	50.629596	-3.454630	South West

Table 2.2: Rothamsted suction-trap sites and their region (See Fig. 2.4 for a map of regions).

2.4.6 Relationship between number of pollen beetles caught in spring and summer

Each calendar year contains two generations of *B. aeneus* – an overwintered spring generation (Overwintered Beetles) and their offspring, which fly in late summer (Newly Emerged Beetles). In order to investigate the relationship between the abundance of these generations, daily data from all 12 suction-traps were used. Count data for each year was divided into the two generations, with an ecologically relevant estimated cut-off date marking the division set at 31st

May/1st June and summed to give a value for each generation in each trap year. Two correlation analyses were carried out 1) comparing the Overwintered Beetles with the following Newly Emerged Beetles and 2) comparing the Newly Emerged Beetles with the following year's Overwintered Beetles. In effect asking whether a) the population size of the parent generation is related to the abundance of their offspring and b) whether the abundance of the emergent generation is reflected in the size of the population post-hibernation. Pearson correlation analyses were performed using the *cor* function in base R and plotted using the *ggscatter* function from the ggpubr library.

The relationship between the overwintered generation in spring and the previous year's new generation in late summer and selected meteorological parameters was investigated more closely using GAMs. For this, data from the two long-term suction-traps (Rothamsted and Broom's Barn) were used. The following meteorological parameters were selected: *Winter Rainfall*, designated as the mean rainfall during the months December, January and February; *Winter Grass Temperature*, designated as the accumulated day degrees of the Grass Minimum Temperature below 0°C during the months December, January and February; *Autumn Rainfall*, designated as the mean rainfall during the months September, October and November; and *Spring Temperature*, designated as the accumulated day degrees of air temperature over 6°C during the months March, April and May. The rainfall parameters were selected to explore the effects of precipitation during overwintering (Winter Rainfall); and waterlogging of the soil prior to, or during, the beetles selecting their overwintering sites (Autumn Rainfall). The temperature parameters were selected to explore the effects of low temperatures at ground level where beetles are overwintering (Winter Temperature) and the direct effects of warm, or cold springs on the emerging overwintered generation in spring (Spring Temperature). Due to the lack of replication – only two sites were used

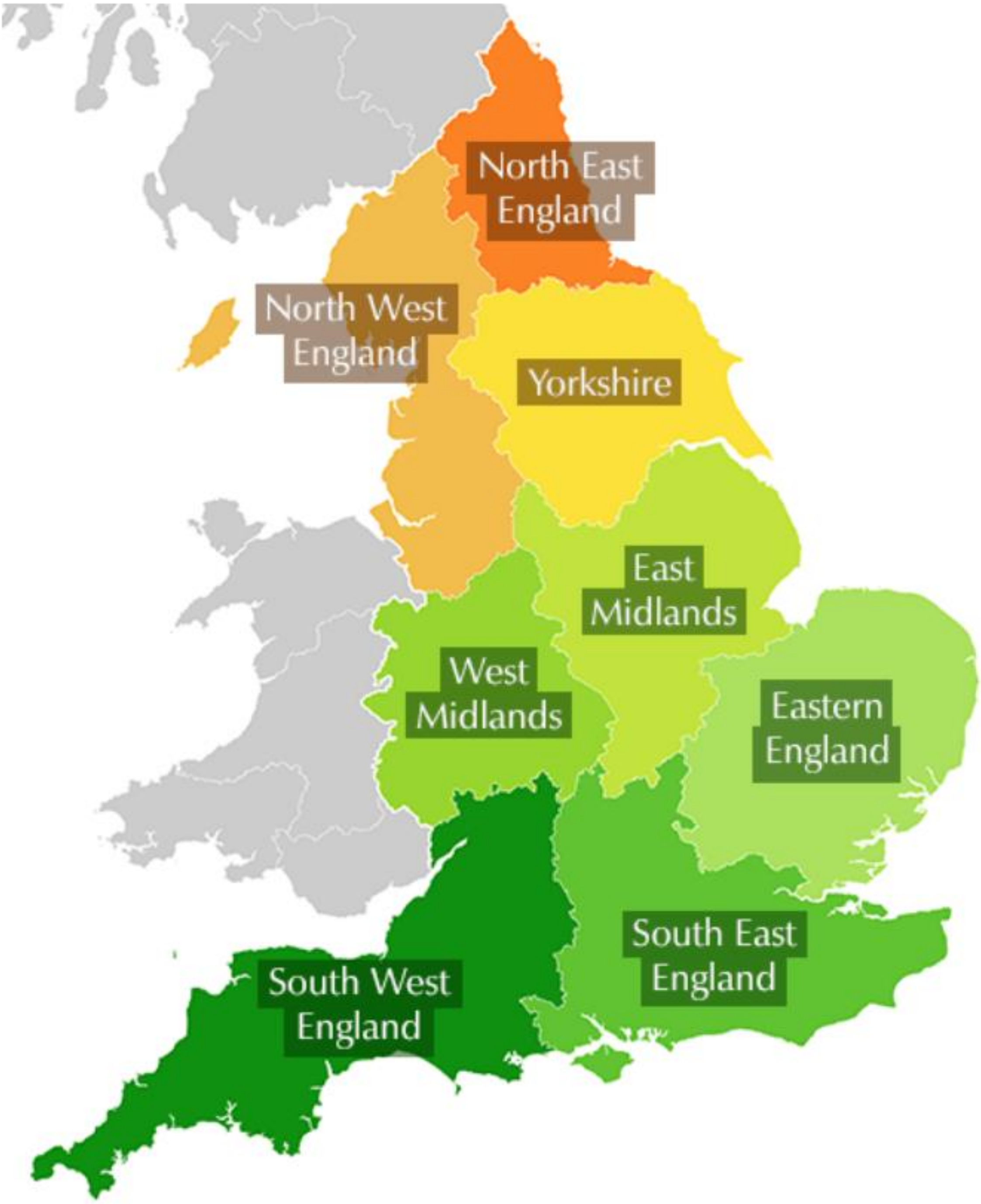


Fig. 2.4: English regions as used in this study (from https://techspolondon.co.uk/england-regions/).

– it was not possible to include any random effects of, for example, year or site. Year was included as a fixed effect. The relevant code is included in Appendix A, Chapter 2.

2.4.7 Drivers of overwintered beetle abundance

To investigate drivers of abundance in the Spring population the tree package (Ripley 2019) was used to fit a regression tree to the data used in 2.4.6 above. This tree is "grown" by binary recursive partitioning; the explanatory parameter variable data are successfully split along their coordinate axes so that, at any node (or branching point), the split is chosen that maximally distinguishes the observation data (response) in the respective branches – at each node a higher value in the explanatory parameter gives a response in the right-hand branch, with a lower value the response is in the left-hand branch, the objective being to find the smallest, most parsimonious partitioning of the data (Brieman *et al*. 1984, Ripley 1996, De'ath & Fabricius 2000). See Appendix A, Chapter 2 for the relevant code.

2.5 Results

2.5.1 Oilseed rape cultivation trend analysis

As Fig. 2.1 showed, oilseed rape cultivation over time has had a positive trend in the UK from 1983 to 2018. This is also true over the shorter-term in England (Fig. 2.5) and south-east England (Fig. 2.6) and, although yield has fluctuated (Fig. 2.7a), production has closely matched area (Figs. 2.6, 2.7b). The smoothed model (Fig. 2.8) suggests that in England the area of OSR has increased between 83% and 142% and that areas sown increased significantly between 2001 and

2011 before levelling out, reaffirming DEFRA reports that production has plateaued since 2012 (DEFRA 2018, 2019a).

Interestingly there is no sign of a significant change of shape in the trend around 2010-2012 suggesting a slow levelling out, which is reflected in the shape of the plot. The indication of significant changes in curvature at the start and end of the series are potentially an artefact of end point influence. Table 2.3 shows the area sown, yield and production data for all regions used in the poptrends analysis.

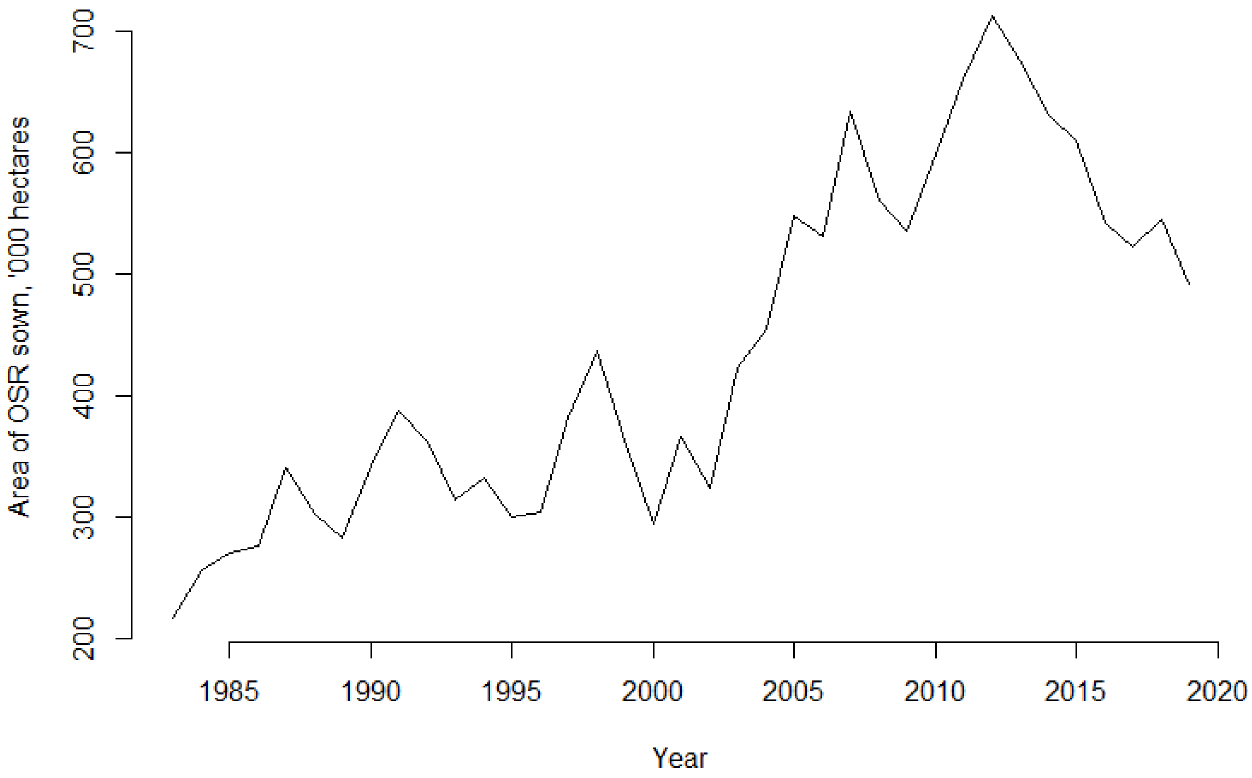


Fig. 2.5: Area of oilseed rape (OSR) sown in the UK 1984-2019, data from DEFRA.

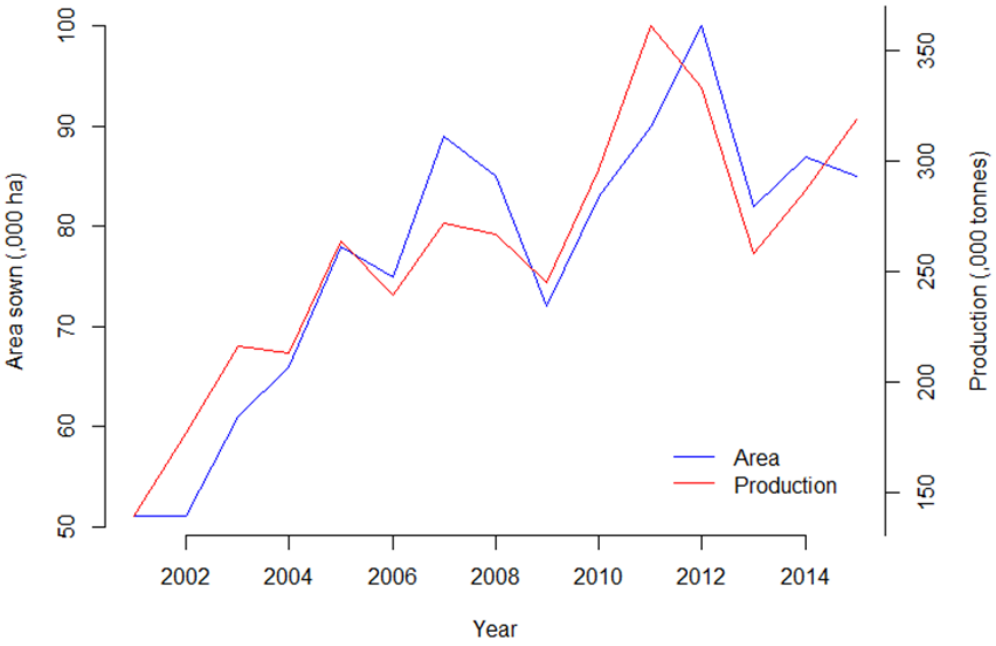


Fig. 2.6: Area of winter oilseed rape sown in the South East of England 2001-2014 (blue) and recorded production for the region (red), data from DEFRA.

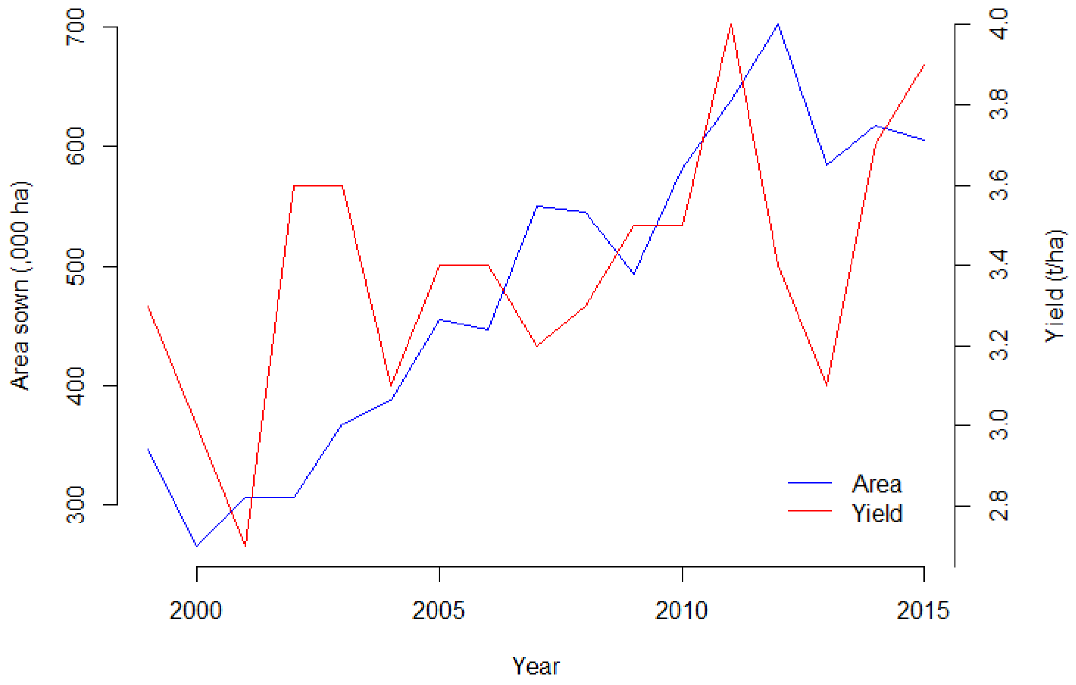


Fig. 2.7a: Area of winter oilseed rape sown in England 1999-2015 (blue) with the recorded yield (red), data from DEFRA (see section 2.3.2).

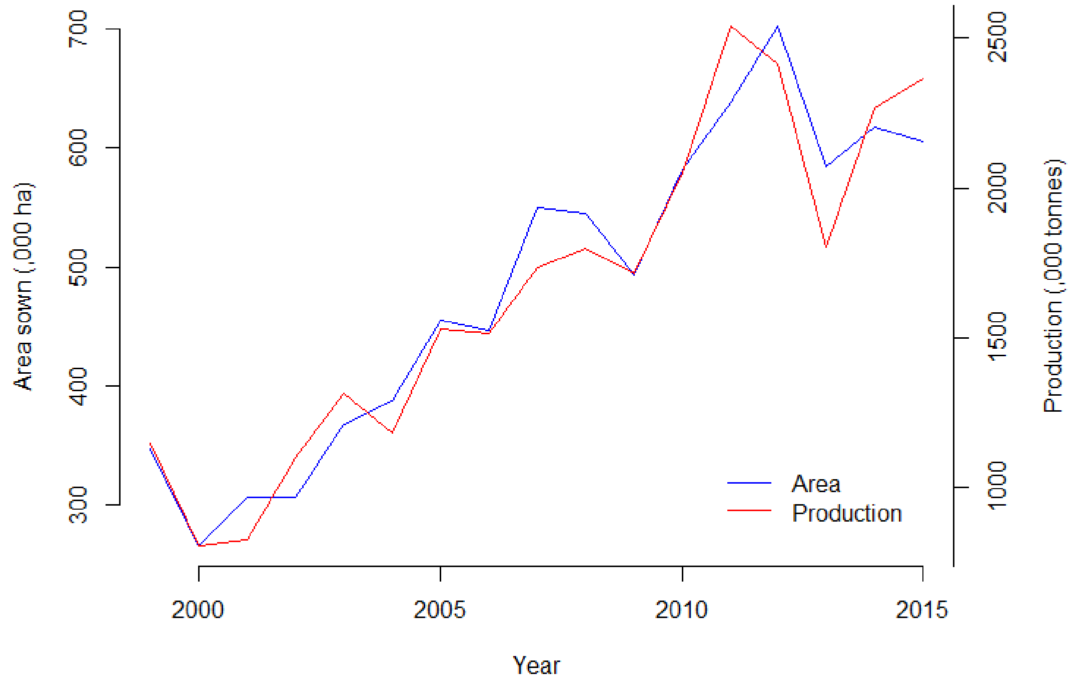


Fig. 2.7b: Area of winter oilseed rape sown in England 1999- 2015 (blue) with the recorded production (red), data from DEFRA (see section 2.3.2).

	England			North East			North West			Yorkshire			East Midlands			West Midlands			Eastern Area			South East			South West		
Year	Area Sown	Yield	Production (tonnes)	Area Sown	Yield	Production (tonnes)	Area Sown	Yield	Production (tonnes)	Area Sown (,000ha)	Yield	Production (tonnes)	Area Sown (,000ha)	Yield	Production (tonnes)	Area Sown (,000ha)	Yield	Production (tonnes)	Area Sown (,000ha)	Yield	Production (tonnes)	Area Sown (,000ha)	Yield	Production (tonnes)	Area Sown (,000ha)	Yield	Production (tonnes)
1999	347	3.3	1149		3.1			2.8			3.5			3.4			3.1			3.5			3.2			2.8	
2000	265	3	805		3.1			3.1			3			3			3			3.1			3			3	
2001	306	2.7	828	16	2.7	42	3	2.6	8	38	2.9	111	78	2.5	197	22	2.8	61	77	2.6	204	51	2.8	139	22	3	66
2002	306	3.6	1099	17	3.7	63	3	3.2	9	39	3.9	151	79	3.6	284	21	4.1	85	73	3.4	249	51	3.5	177	23	3.5	81
2003	367	3.6	1312	19	3.6	68	3	3.6	9	46	3.9	178	91	3.5	320	28	3.6	98	87	3.5	304	61	3.5	216	32	3.8	119
2004	387	3.1	1183	22	3.2	69	3	3.4	10	56	3.5	196	95	3.1	293	29	3	89	79	2.7	212	66	3.2	213	37	2.7	100
2005	455	3.4	1527	23	3	69	4	3.7	13	62	3.6	227	112	3.2	362	35	3.1	109	97	3.5	342	78	3.4	264	43	3.2	139
2006	447	3.4	1512	22	3.2	71	3	3.6	12	60	3.7	219	111	3.2	360	32	3.5	113	102	3.6	361	75	3.2	239	41	3.3	137
2007	550	3.2	1734	26	3.6	94	4	3.6	14	75	3.3	250	136	3	402	41	3.3	134	128	3.1	395	89	3	272	50	3.5	173
2008	545	3.3	1799	26	3.2	84	4	3.3	13	77	3.3	251	135	3.4	459	41	3.5	146	128	3.3	418	85	3.1	267	48	3.3	161
2009	493	3.5	1716	19	2.9	54	4	3.5	13	69	3.5	239	124	3.4	423	38	3.7	141	123	3.6	440	72	3.4	245	44	3.7	160
2010	582	3.5	2052	24	3.2	79	5	3.7	17	84	3.6	306	148	3.6	528	47	3.6	167	139	3.4	477	83	3.6	296	52	3.5	181
2011	638	4	2537	27	4	107	5	3.8	21	90	4.3	390	164	3.9	641	53	3.9	207	151	3.7	559	90	4	361	57	4.4	252
2012	702	3.4	2411	29	2.9	85	6	3.3	19	98	3.5	340	180	3.4	615	60	3.3	200	166	3.7	608	100	3.3	333	63	3.3	211
2013	584	3.1	1800	21	2.7	56	4	2.7	11	79	3	234	146	2.9	429	48	3	145	150	3.4	504	82	3.1	258	53	3	161
2014	618	3.7	2266	24	3.7	87	5	3.9	18	83	3.8	318	155	3.7	578	50	3.9	196	158	3.7	590	87	3.3	287	58	3.3	190
2015	605	3.9	2363	24	4.3	104	5	3.6	18	83	4.2	351	154	4.1	629	48	4.1	197	151	3.6	539	85	3.8	319	56	3.7	207

Table 2.3: Winter oilseed rape statistics (area sown, yield and production) for England and its regions 1999-2015 data from DEFRA (see section 2.3.2).

2.5.2 Long-term beetle data

In total 37,143 beetles were recorded. Table 2.4 details the annual counts recorded at each site. Both Rothamsted and Broom's Barn showed similar values for *B. aeneus* abundance in each year (Fig. 2.9), and indeed they were closely correlated (Fig. 2.10) (t = 6.549, df = 28, p <0.001, R = 0.78). Trend analysis using the poptrends package suggests that beetle populations at the two sites have increased by 162% (s.e. 5.5%, 607%); they showed no significant trend in numbers until around 2006, when a significant upturning began a positive trend that continued to increase significantly until the dataset ends in 2018 (Fig. 2.11). At the seasonal level there was a non-significant trend in abundance in all years except around 2010-12 where the upward trend was significant in the spring (Fig. 2.12) whilst the trend in the summer matched the annual trend, although there was no significant change in the curvature (Fig. 2.13). At the monthly level, populations in March/April showed no significant trend (Fig. 2.14), whilst May, June and August/September showed a similar pattern to the annual data (Figs. 2.15, 2.16 & 2.18) although the magnitude of the trend differs across months. March/April and May showed no significant curvature changes outside the end points (Figs. 2.14, 2.15) with the upward trend in May starting later than the full dataset, around 2007 (Fig. 2.15). June and August/September showed upward curve changes and 2006-7 and 2009-10 (Figs. 2.16 & 2.18), also comparable to the full dataset though the significant positive trend started earlier, around 2004. July showed a more complex result, with upward curvature noted around 1995-97, 2001 and 2005-6 but only the trace of a significantly increasing trend around 2011 and a significant decline in 2003-4 (Fig. 2.17). Of those that show a significant upward trend the magnitude is similar across all datasets with the exception of August/September which has much higher trend.

Year	Newcastle	York	Preston	Hereford	Wellesbourne	Kirton	Broom's Barn	Rothamsted	Writtle	Silwood Park	Wye	Starcross
1987								388				
1988								435				
1989								1508				
1990								1126				
1991								148				
1992								107				
1993								365				
1994								785				
1995								787				
1996								481				
1997								137				
1998								156				
1999								208				
2000								372				
2001								222				
2002								406				
2003								267				
2004								474				
2005							233	31				
2006							141	165				
2007							142	182				
2008							159	194				
2009							436	274				
2010							1189	928				
2011							964	508				
2012							953	372				
2013							714	283				
2014							1838	784				
2015	526	409	16	25	686	714	2495	1488	1885	140	2056	229
2016	729	668		84	628	670	2864	1149	1759	270	1479	295
2017	1267	1132		107	788	1072	1544	1091	836	171	1063	417
2018	708	308		181	678	584	1534	685	648	148	930	165

Table 2.4. Number of *Brassicogethes aeneus* recorded each year from a network of 12 suction-traps operating across the UK. Recording started in 2015 for all sites except for Rothamsted and Brooms Barn which were recorded from 1987 and 1989, respectively. Blank cells indicate that the trap was not investigated for that year.

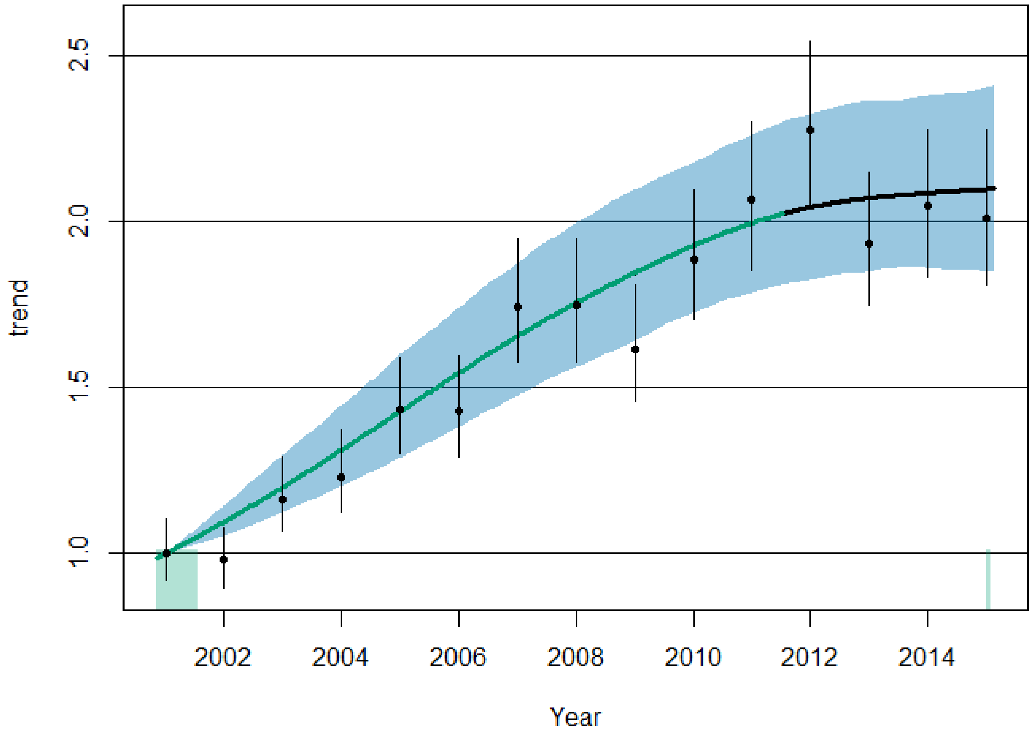


Fig. 2.8: Standardized trends for area of winter oilseed rape in England with random effects (dots and whiskers), indicating yearly mean and variances, and 95% confidence intervals (blue). All trends are relative measures that are standardised against the total predicted area in the first year (the reference year). Significantly different increasing short-term trends at the 5% level are coloured green and extend for the time period of the increasing trend. These short-term trends are imposed on top of the long-term trend, coloured in black. The shape of change (i.e. second derivatives) are indicated along the x axis, with upturning (∪ - green bars) curvatures in the trend indicated.

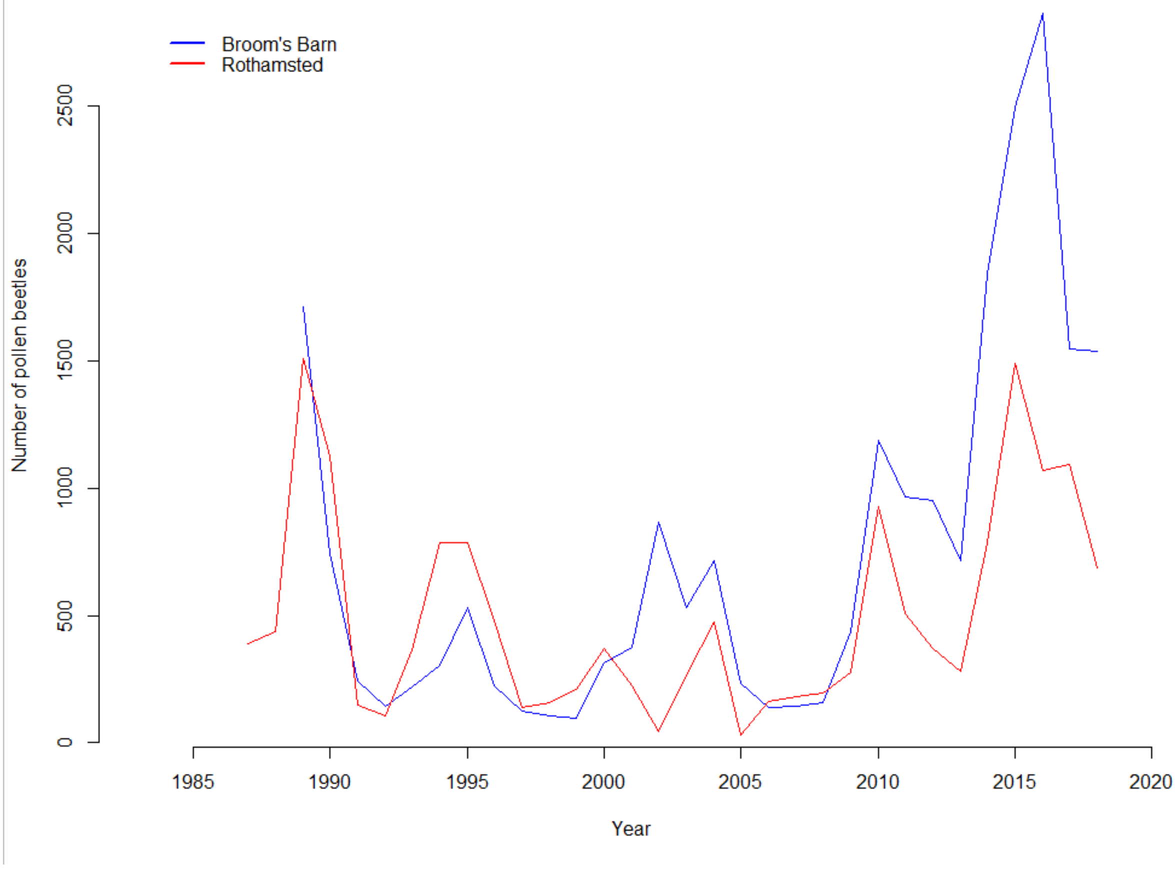


Fig. 2.9: Annual count of *Brassicogethes aeneus* caught at Rothamsted and Broom's Barn suction-trap sites 1987-2018.

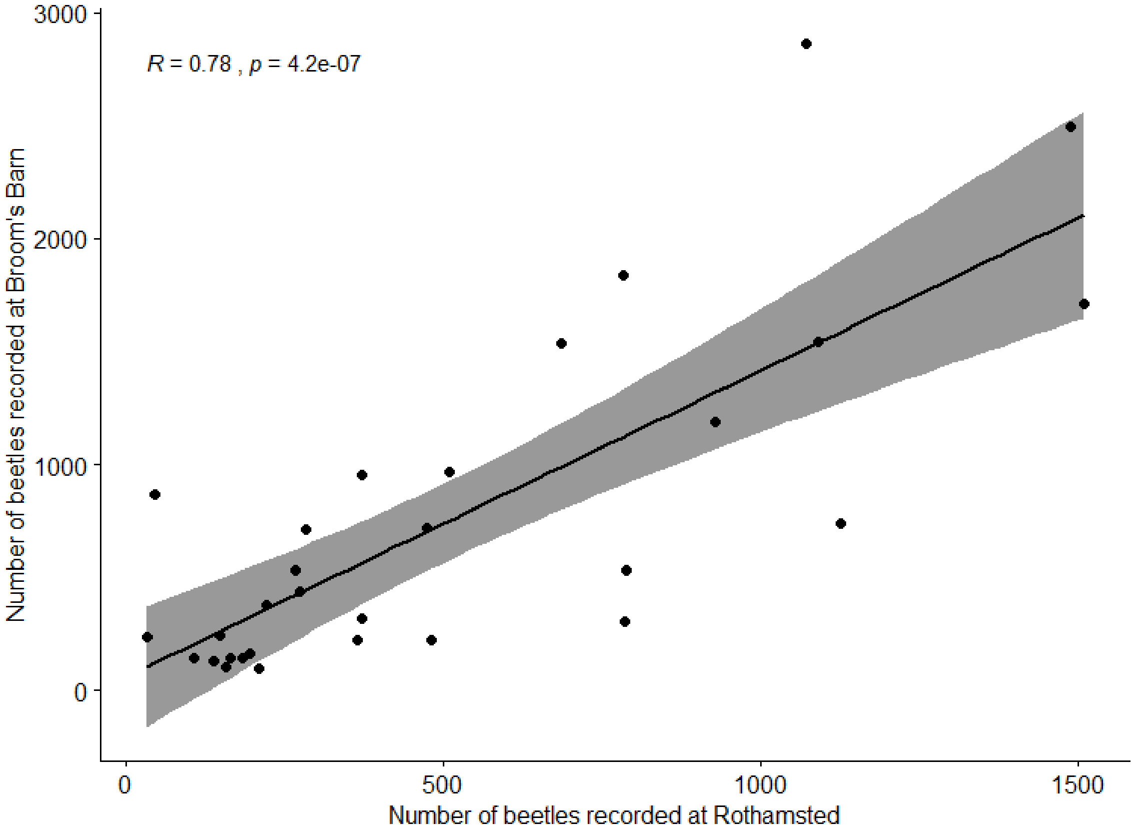


Fig. 2.10: Correlation between annual counts of *Brassicogethes aeneus* caught in suction-traps at Rothamsted and Broom's Barn 1987-2018.

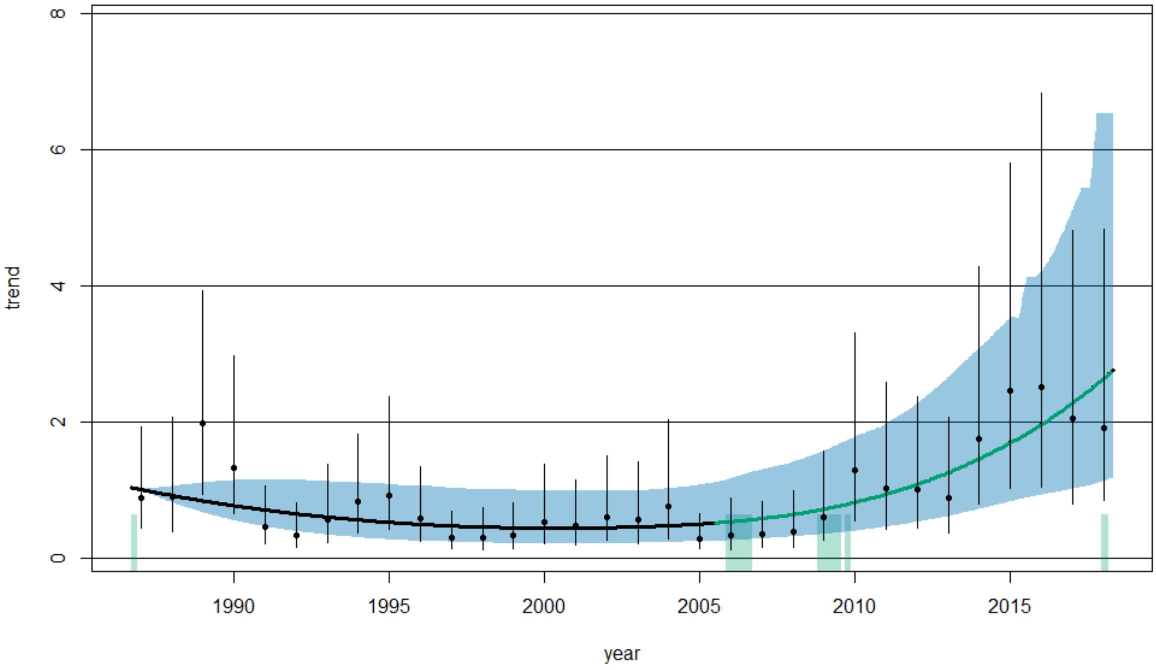


Fig. 2.11: Standardized regional trends for the annual count of *Brassicogethes aeneus* caught in suction-traps at Rothamsted and Broom's Barn. All trends are relative measures that are standardised against the total predicted area in the first year (the reference year). See Fig. 2.8 for a detailed explanation of the features of this plot.

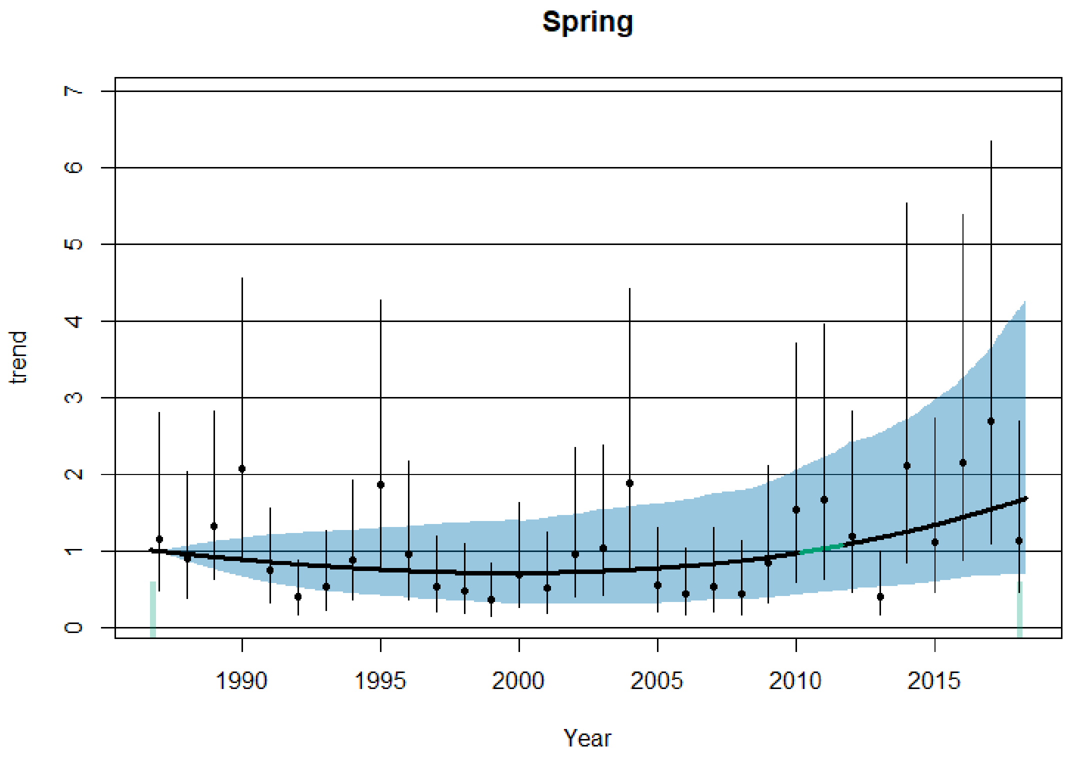


Fig. 2.12: Population trend of *Brassicogethes aeneus* in spring months (up to June 1st) at Rothamsted and Broom's Barn. All trends are relative measures that are standardised against the total predicted area in the first year (the reference year). See Fig. 2.8 for a detailed explanation of the features of this plot.

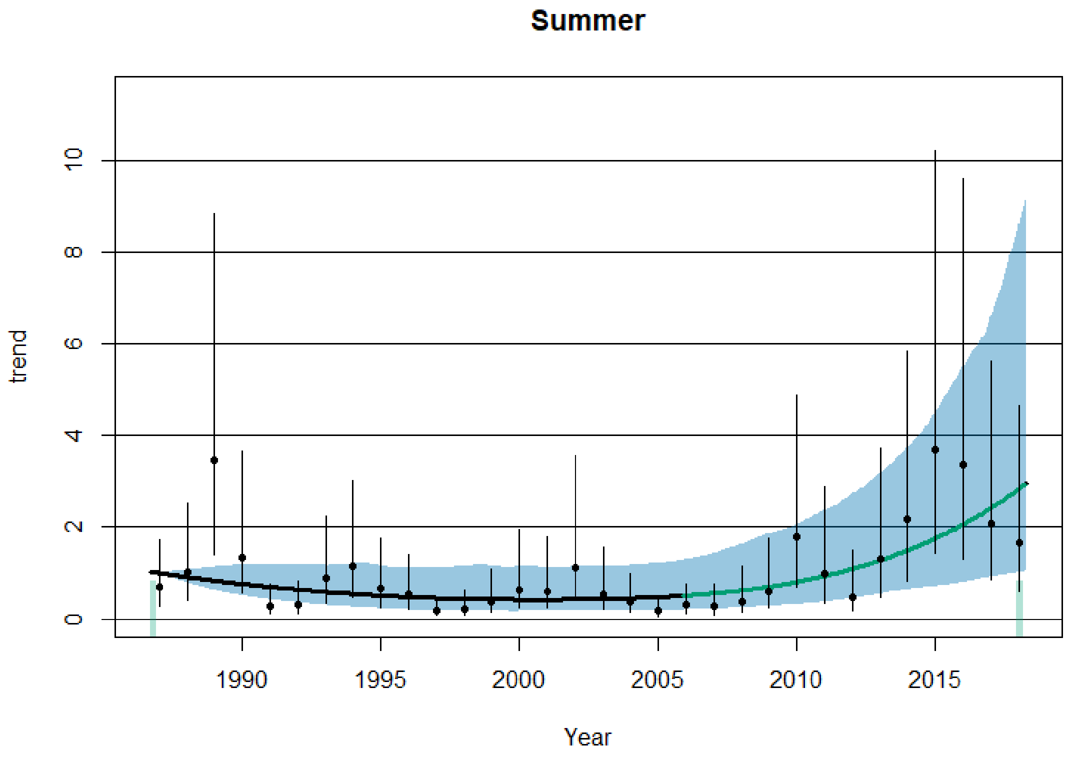
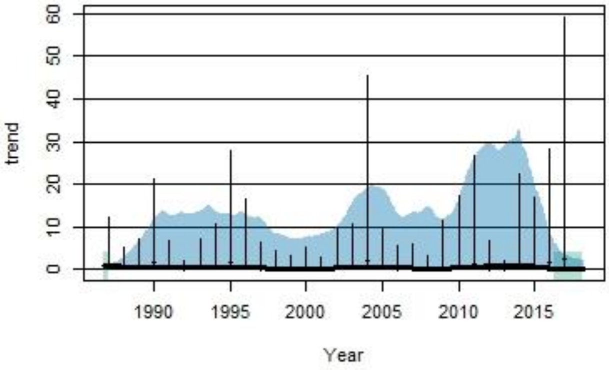
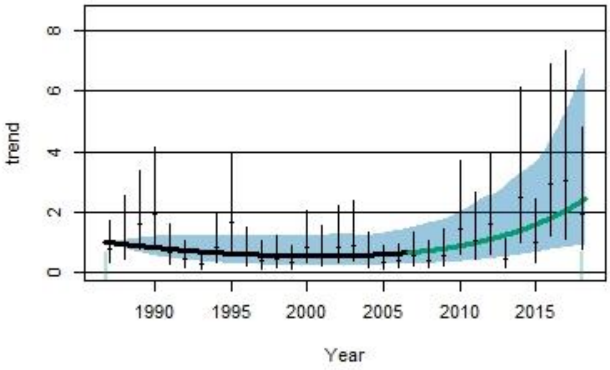


Fig. 2.13: Population trend of *Brassicogethes aeneus* in summer months (June 1st onwards) at Rothamsted and Broom's Barn. All trends are relative measures that are standardised against the total predicted area in the first year (the reference year). See Fig. 2.8 for a detailed explanation of the features of this plot.

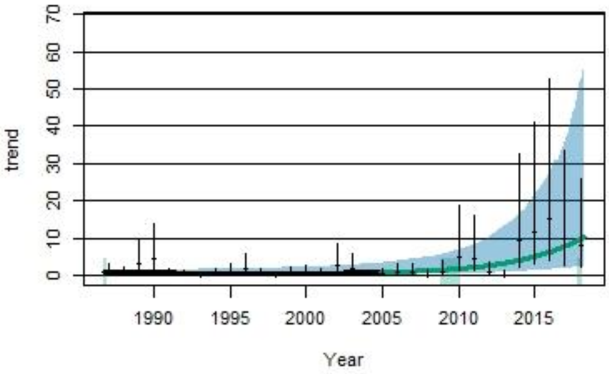
2.14: March and April



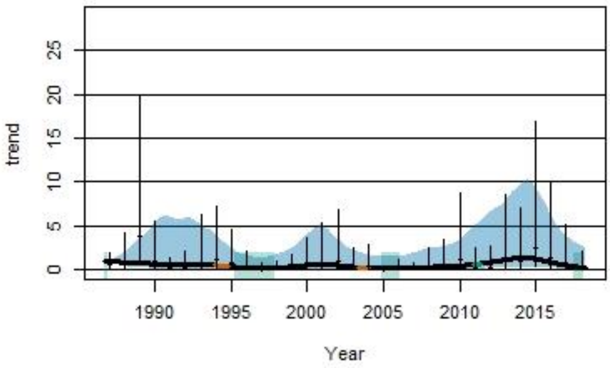
2.15: May



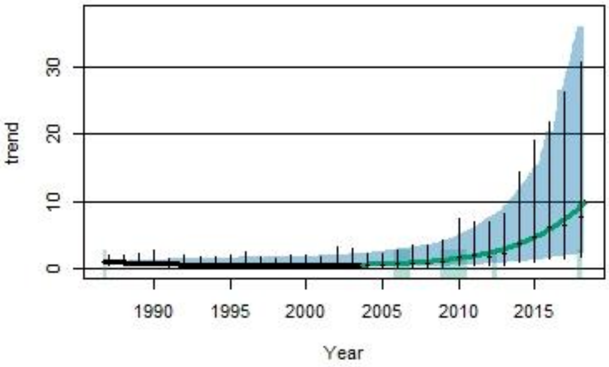
2.16: June



2.17: July



2.18: August and September



Figs 2.14-2.18: Monthly or bi-monthly population trends of *Brassicogethes aeneus* abundance at Rothamsted and Broom's Barn. See Fig. 2.8 All trends are relative measures that are standardised against the total predicted area in the first year (the reference year). for a detailed explanation of the features of these plots.

2.5.3 Relationship between number of pollen beetles and area of oilseed rape at different temporal and spatial scales

The GAM of annual *B. aeneus* counts from suction-traps at Rothamsted and Broom's Barn showed a similar, and significant, relationship with year as revealed in the poptrends analysis. There was a significant negative relationship found between the area of OSR and *B. aeneus* at Broom's Barn (Table 2.5, Fig. 2.19), this relationship was also negative, although non-significant, at Rothamsted (Table 2.5, Fig. 2.20). At a regional scale there was no correlation between numbers at Broom's Barn (Fig. 2.21) (t = 2.0865, df = 13, p-value = 0.057, R = 0.5) and the area of oilseed rape sown in the eastern England, whilst Rothamsted showed a weak positive correlation between counts and area sown (Fig. 2.22) (t = 2.205, df = 13, $p < 0.05$, R = 0.52). The annual number of beetles in 2015 was weakly correlated with the area of winter oilseed rape crop in the respective region (Fig. 2.23) (t = 2.3924, df = 10, $p < 0.05$, R = 0.6).

	Broom's Barn				**Rothamsted**			
A. parametric coefficients	Estimate	Std. Error	t-value	p-value	Estimate	Std. Error	t-value	p-value
(Intercept)	6.191	0.06	103.501	< 0.0001	6.072	0.12	50.806	< 0.0001
B. smooth terms	edf	Ref.df	F-value	p-value	edf	Ref.df	F-value	p-value
s(Area)	1.968	9	13.139	0.0001	0.874	9	1.592	0.108
s(Year)	7.866	9	191.46	< 0.0001	2.525	9	17.496	< 0.0001

Table 2.5: Summary statistics for GAM of the relationship between *Brassicogethes aeneus* abundance at Rothamsted and Broom's Barn and year and area of oilseed rape.

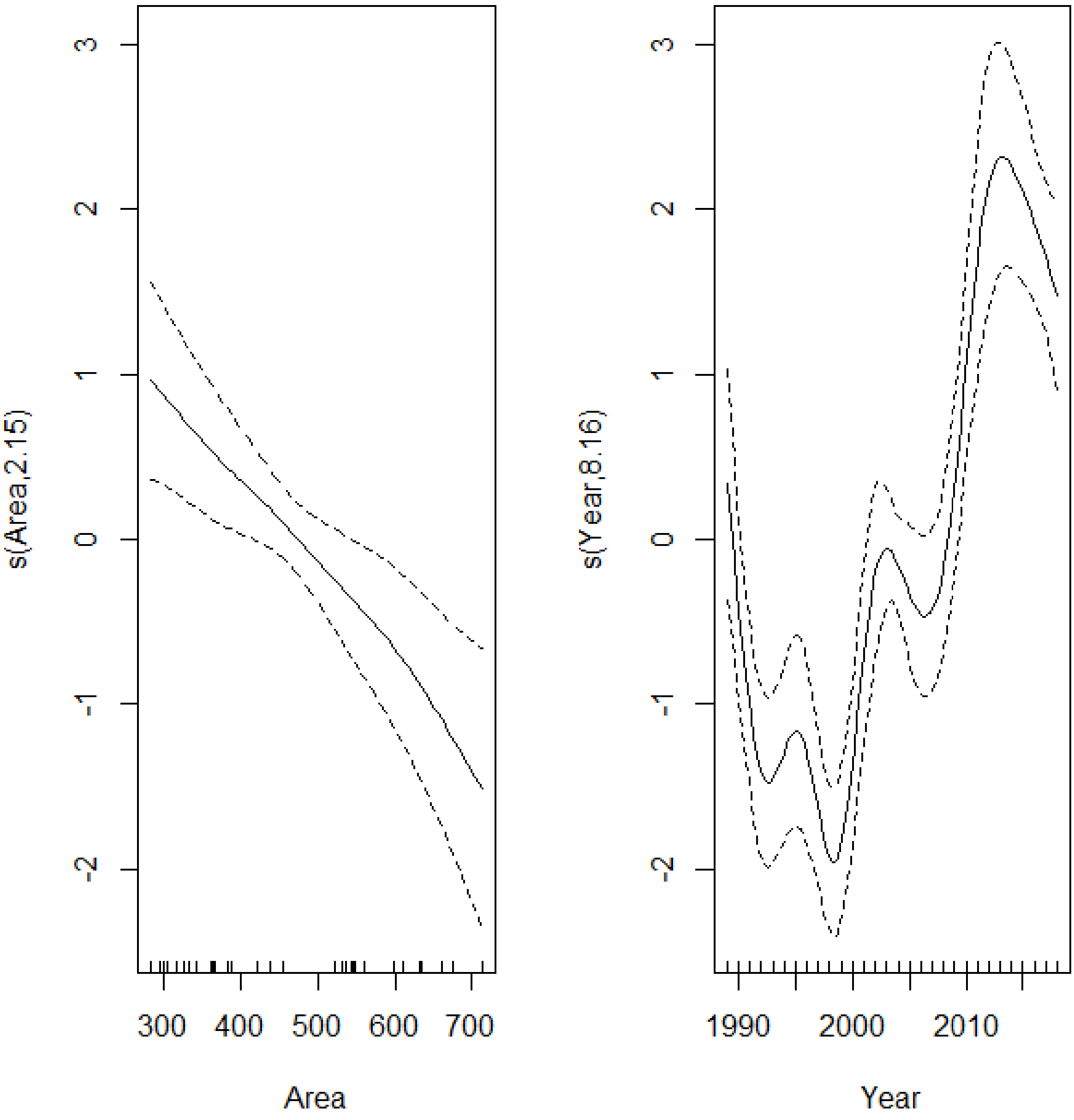


Fig. 2.19: GAM of the relationship between the area (left) of oilseed rape sown in the UK and year (right) with annual *Brassicogethes aeneus* numbers caught at Broom's Barn 1989-2018. The estimated smoothed terms are a transformed function of area or year which on the *y* - axis is centred on zero and scaled by the effective degrees of freedom. The solid line shows the calculated relationship with 95% confidence intervals represented by dashed lines. Rug marks on the x axis represent sampling points. Counts of *B. aeneus* are negatively related to area of crop (left) and show a complex relationship with year (right).

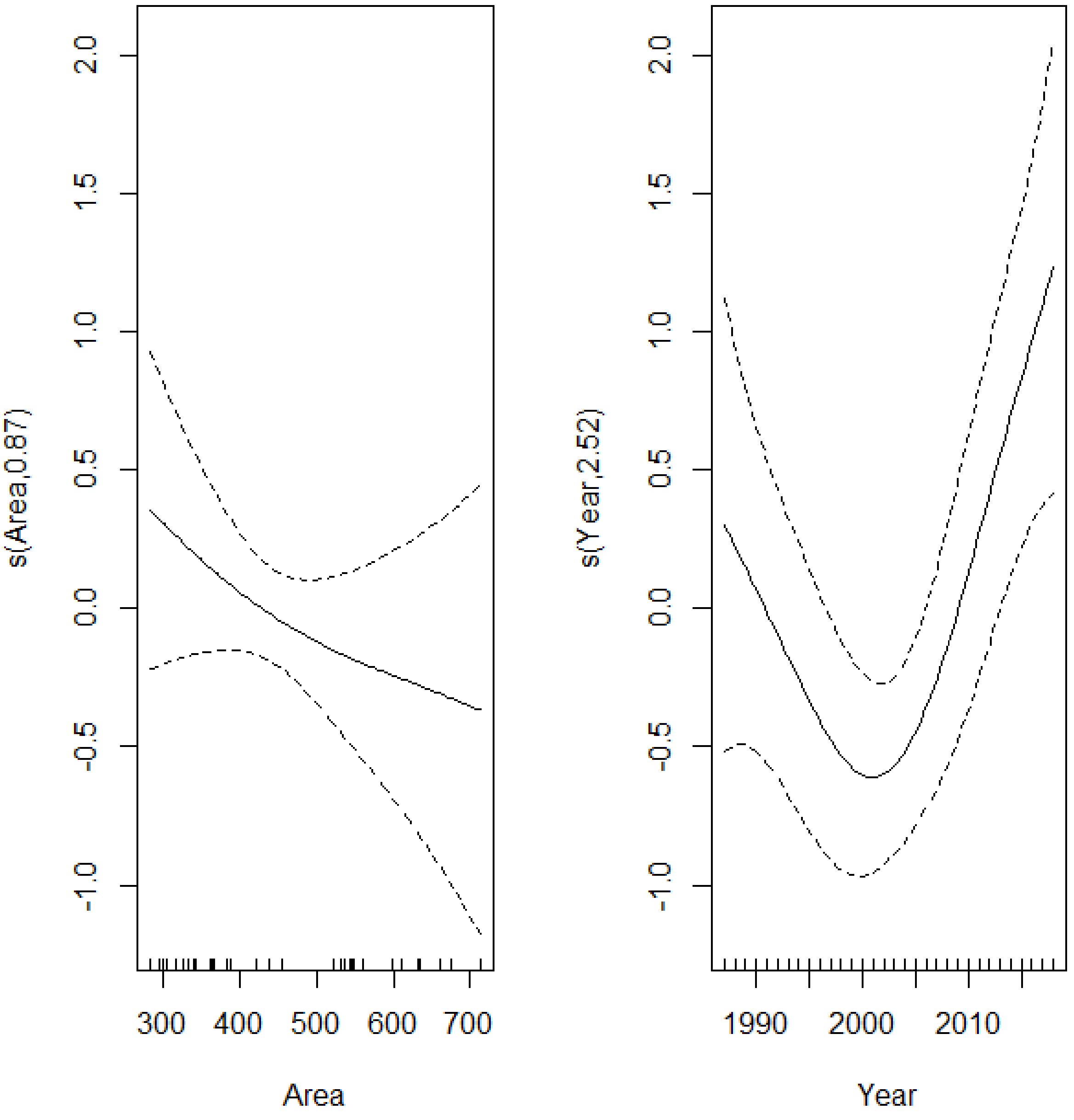


Fig. 2.20: GAM of the relationship between the area (left) of oilseed rape sown in the UK and year (right) with annual *Brassicogethes aeneus* numbers caught at Rothamsted 1987-2018. The estimated smoothed terms are a transformed function of area or year which on the *y* - axis is centred on zero and scaled by the effective degrees of freedom. The solid line shows the calculated relationship with 95% confidence intervals represented by dashed lines. Rug marks on the x axis represent sampling points. Counts of *B. aeneus* are negatively related to area of crop (left) and show a u-shape relationship with year (right).

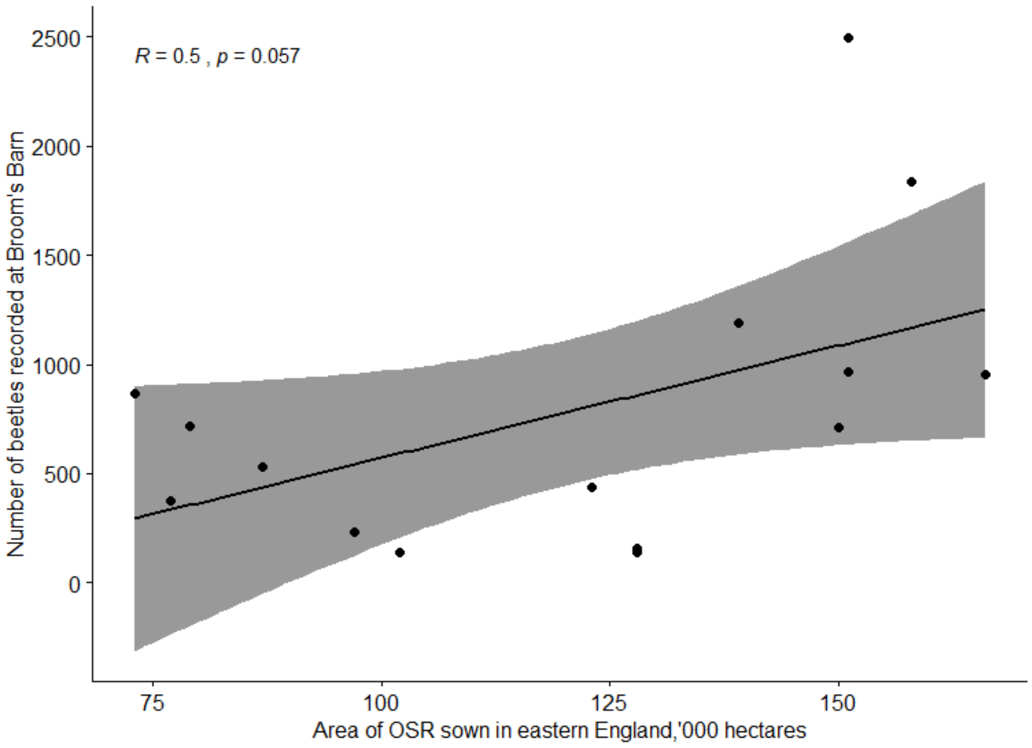


Fig. 2.21: Relationship between area of oilseed rape (OSR) sown in eastern England and the annual count of *Brassicogethes aeneus* in the suction-trap at Broom's Barn 2001-2015.

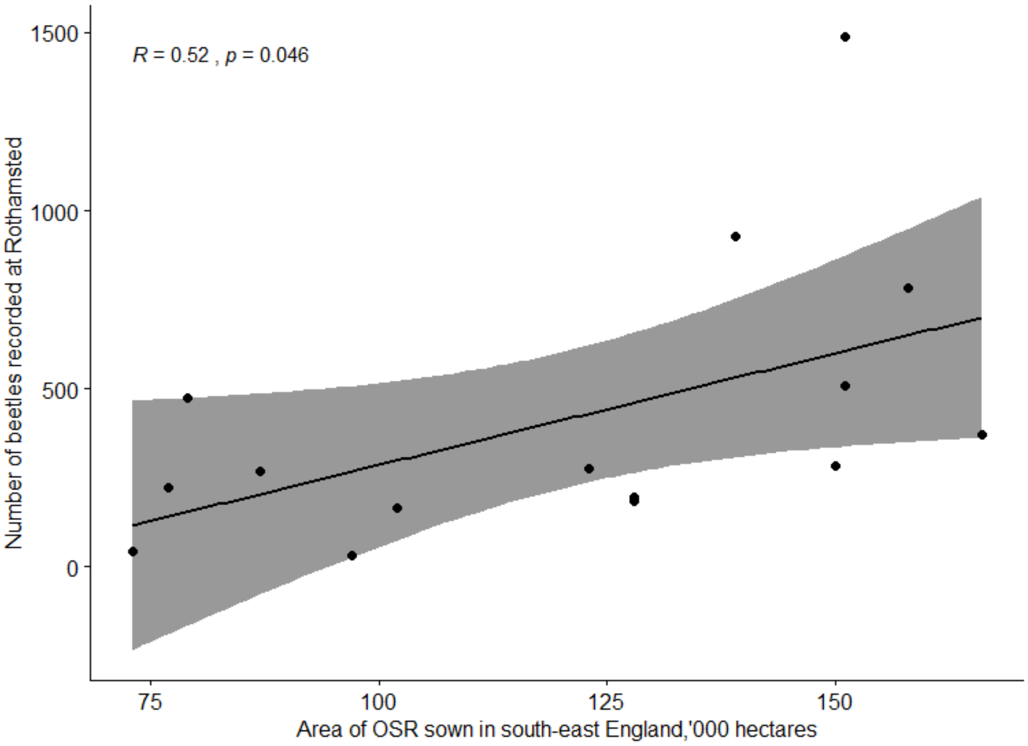


Fig. 2.22: Relationship between area of oilseed rape (OSR) sown in eastern England and the annual count of *Brassicogethes aeneus* in the suction-trap at Rothamsted 2001-2015.

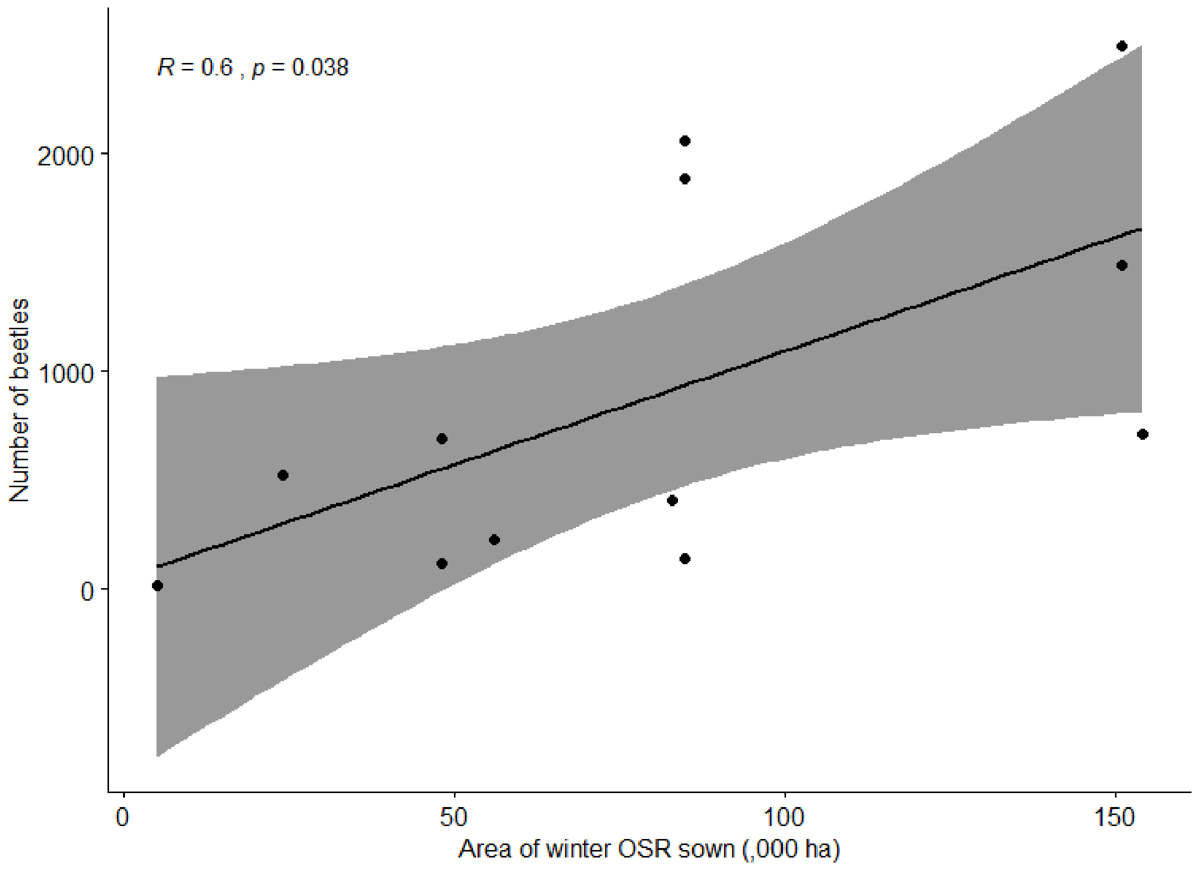


Fig. 2.23: Relationship between number of *Brassicogethes aeneus* caught in the Rothamsted Insect Survey suction-trap network and the area of winter oilseed rape (OSR) in their respective regions in 2015 (see Table 2.2).

2.5.4 Relationship between spring and summer counts of pollen beetles at all sites

Correlation analysis suggested that both the relationship between Overwintered Beetles and the following Newly Emerged Beetles was significantly positive (Fig. 2.24) (t = 5.2832, df = 96, p < 0.001), as was the relationship between the Newly Emerged Beetles entering hibernation and the Overwintered Beetles emerging in the spring (Fig. 2.25) t = 11.108, df = 85, p < 0.001). There is an interesting pattern of scatter in the first plot, with Overwintered Beetles numbers in the low 200s leading to high Newly Emerged Beetles counts at Hereford, Wellesbourne and Broom’s Barn.

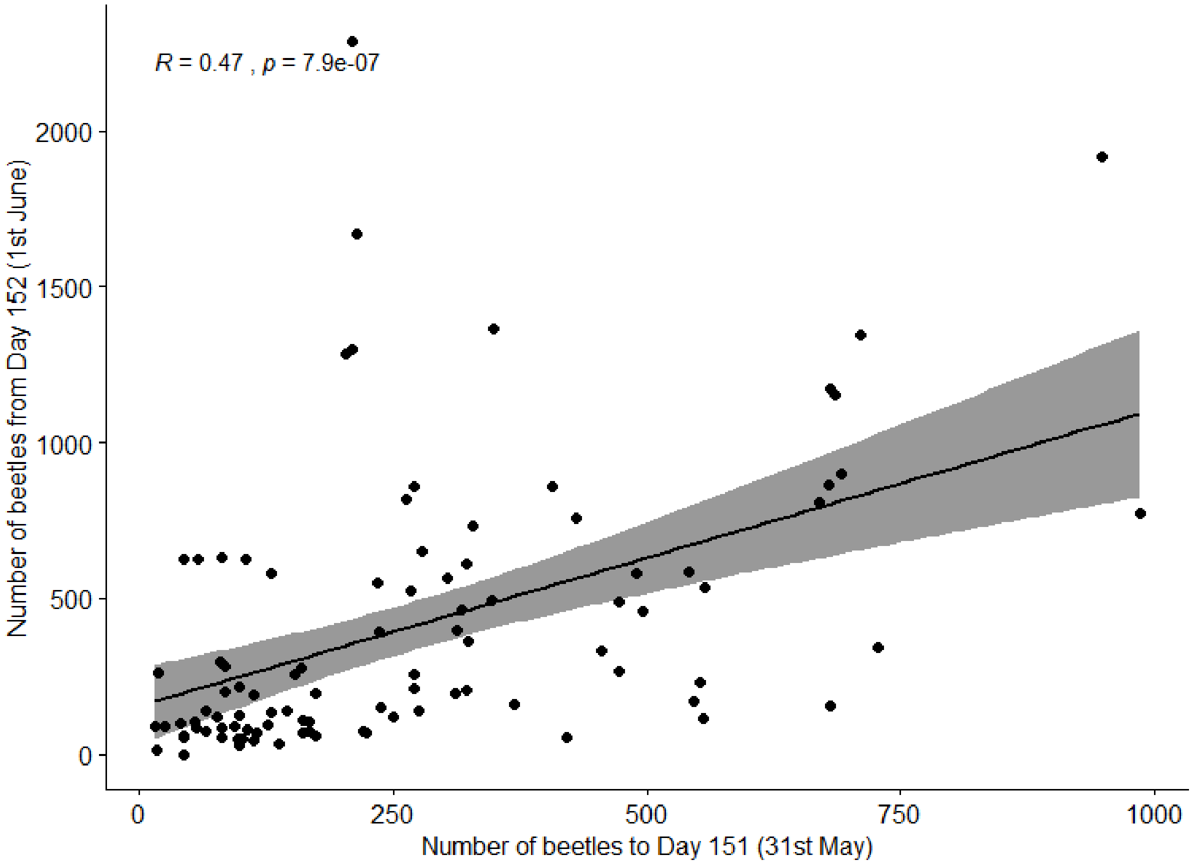


Fig. 2.24: Correlation plot showing relationship between the overwintered generation of *Brassicogethes aeneus* in spring (beetles caught before June 1st) and the subsequent newly emerged generation in summer (beetles caught after June 1st).

Closer examination of the potential drivers of the overwintered generation population size in spring resulted in a model including all the weather parameters selected. Despite several terms being non-significant in the model (Table 2.6), model selection using AIC suggested that this model was the best option compared to models that excluded parameters. GAM plots (Fig. 2.26) showed that year and winter temperature had linear effects on overwintered generation numbers, with

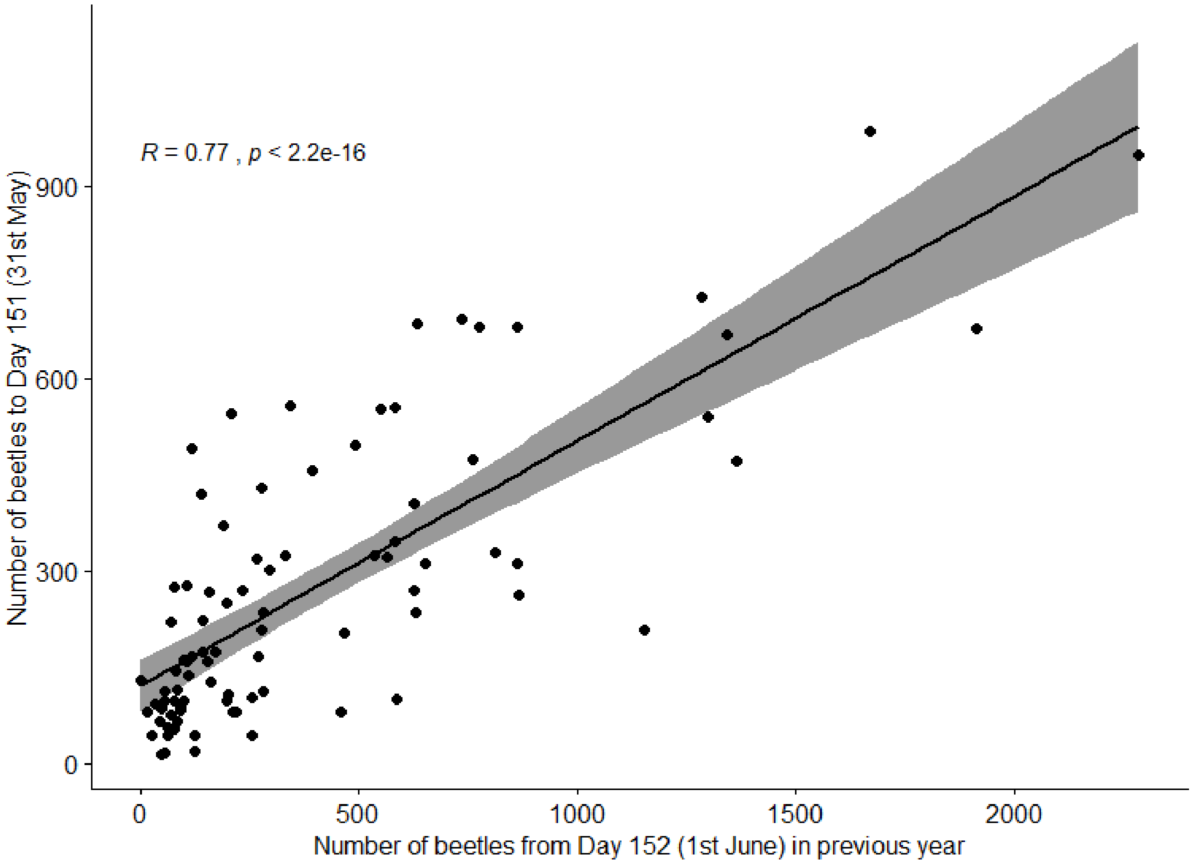


Fig. 2.25: Relationship between the emergent new generation of *Brassicogethes aeneus* in summer (beetles caught after June 1st) and the subsequent overwintered (spring) generation (beetles caught before June 1st the following year).

later years and colder temperatures having positive effects on beetle population numbers. The relationship with other parameters was more complex: winter rainfall had little effect though extremely wet winters appear to result in higher beetle counts in spring (although it should be noted that this is based on very few data points – only three above a mean daily rainfall of 3mm); increased autumn rainfall appears to have a generally negative effect on beetle counts at the extremes, although between ~1.5mm/day and 3mm/day the effect of greater rainfall is positive;

spring temperature has a humped relationship with beetle counts, with extreme high and low temperatures having negative influence on numbers. The size of the new generation population in summer had a positive effect on overwintered beetle counts in spring, though not linear, increasing Newly Emerged Beetles abundance generally led to an increase in Overwintered Beetles counts, however it must be noted that beyond ~750 beetles the data are sparse, which may explain the levelling of the relationship between 500 & 1000 Newly Emerged Beetles counts.

The tree plot (Fig. 2.27) indicates that the population of the Newly Emerged Generation (Summer) is the most important factor in explaining the population in the following Spring, with counts from June 1st onwards above 478.5 leading to counts in the following Spring that are higher than if the counts post-June 1st are below this 478.5 threshold (Fig 2.26). Further, at very high counts of beetles (1219+) post-June 1st, a mean number of 673.6 beetles are found in Spring. Between 479 & 1218 Newly Emerged Generation beetles the accumulated grass minimum temperature (below zero) is important, with accumulated winter temperature below -167.55°C leading to higher counts (mean no. = 460.2) and warmer winter temperatures giving lower counts (mean 287.6). Where low Summer beetle counts (130 or fewer) occur, low numbers (mean 87.22) are found the following Spring. Finally, if Summer counts are between 131 and 478 then the Overwintered Generation counts are higher post-2001 (mean 331.1) than before (mean 183.7).

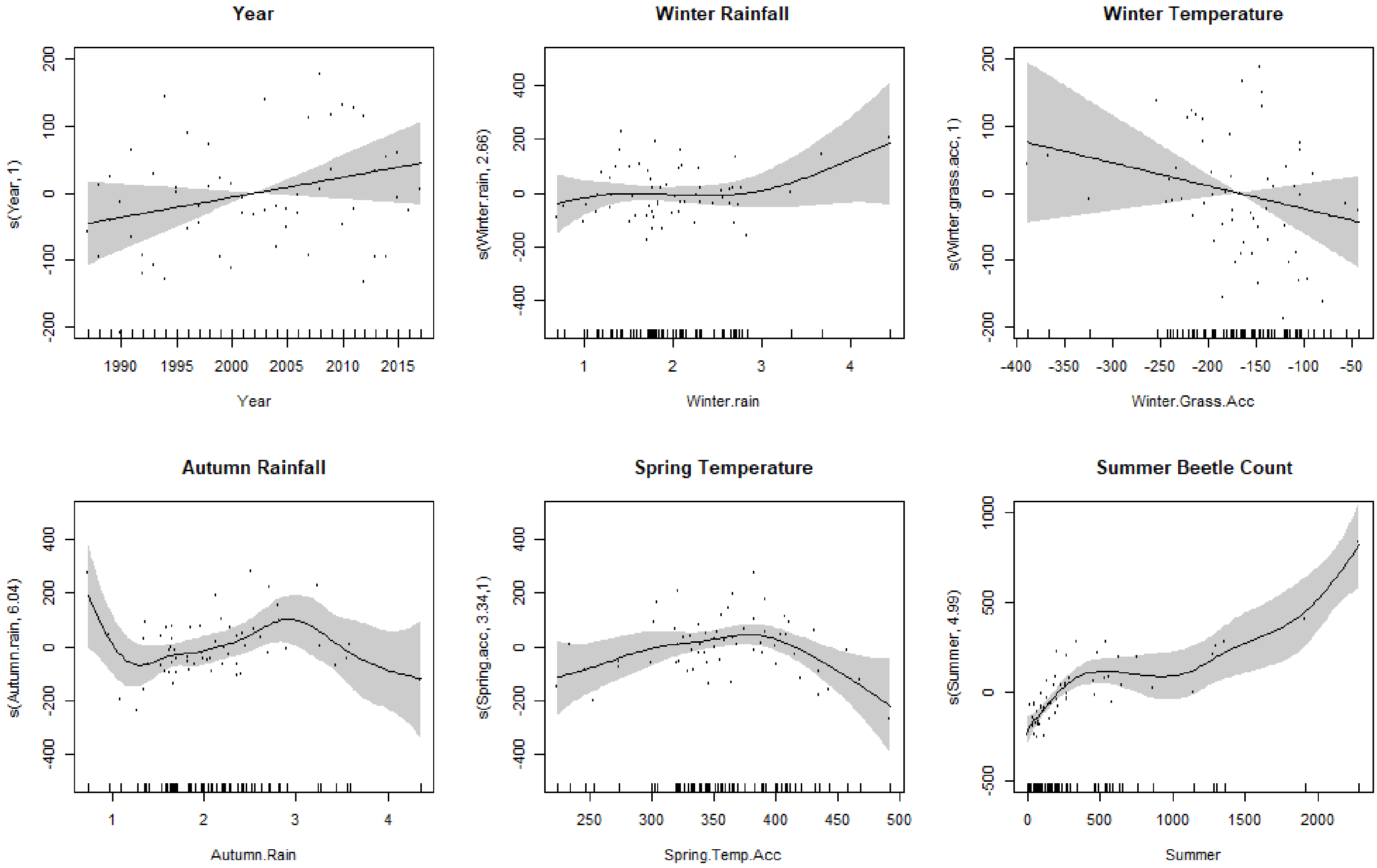


Fig. 2.26: Generalized Additive Models of the relationship between the numbers of *Brassicogethes aeneus* caught in suction-traps in the spring and meteorological and population parameters selected to explain the counts. The estimated smoothed terms are a transformed function which on the *y* - axis is centred on zero and scaled by the effective degrees of freedom. This GAM is estimated by penalized restricted maximum likelihood to minimise overly wiggly smoothers, hence linear terms with estimated degrees of freedom for year and winter temperature equal to 1, as indicated on the *y* axes. Conversely, the remaining smoothers were non-linear, the most complex of which was autumn rainfall. Only accumulated spring temperature and summer beetle numbers were significant.

A. parametric coefficients	**Estimate**	**Std. Error**	**t-value**	**p-value**
(Intercept)	241.3517	13.5578	17.8016	< 0.0001
B. smooth terms	edf	Ref.df	F-value	p-value
s(Year)	1	1	2.1051	0.1545
s(Winter.rain)	2.6575	3.3068	0.917	0.4626
s(Winter.Grass.Acc)	1	1	1.6062	0.2123
s(Autumn.Rain)	6.0412	7.0431	1.9913	0.0771
s(Spring.Temp.Acc)	3.3432	4.1091	2.8991	0.0337
s(Summer)	4.9882	5.8891	17.2944	< 0.0001

Table 2.6: Summary statistics for Generalized Additive Models of relationship between the numbers of *Brassicogethes aeneus* caught in suction-traps in the spring and meteorological and population parameters selected to explain the counts.

2.6 Discussion

2.6.1 Increase of pollen beetles and oilseed rape

It is clear that *B. aeneus* populations are increasing in the UK along with an increase in the area of its cropped host plant. However, there is no correlation in the increase between these variables in the long-term at a national scale (Figs. 2.19, 2.20); regionally, there is at best only a weak positive correlation between 2000 and 2015 (Figs. 2.21-2.23). The contrasting shapes of change derived by the second derivative (continual increase in the *B. aeneus* population (Fig. 2.11) as opposed to a trend reaching asymptote in OSR (Fig. 2.8)), and the fact that *B. aeneus* populations continue to increase beyond 2011 goes some way to explaining the lack of a strong relationship between the two variables in most cases and suggests that *B. aeneus* populations may not yet have reached equilibrium with habitat area, or that the recent development of pyrethroid insecticide resistance in *B. aeneus* has led to increased survival rates. Alternatively, these results could mean that *B. aeneus* numbers are not driven by area of cropped habitat, or that a key driver has not been included in this analysis. The strong negative relationship found between Broom's Barn counts and cropped area in the UK (Fig. 2.19) is hard to explain, although indications are that these migrating beetles are likely to be determined at scales not dependent on the region. It could also reflect the fact that, in later years the area of OSR in the East and South-east regions has stayed stable or has decreased with cropping area increases elsewhere in the country while *B. aeneus* numbers continue to increase. However, in the absence of regional cropping data for the post-2015 period up this can, currently, only be a speculation. Other potential reasons include the presence of a significant lag between increased resource and increased pest populations or management changes in the region. As will be seen in Chapter 4, suction-trapped catches of migrant beetles may not accurately reflect *B. aeneus* counts at the field level and represent a limitation in this study. With cropped area of OSR set to drop significantly in the short term (DEFRA 2020) it will be interesting to see whether *B. aeneus* populations continue to increase or whether they follow the well-established population cycle dynamics seen in other ecological systems and reach a point where numbers crash before returning to a density-dependent multi-year cycle, albeit at a higher abundance than previously (Strong 1979, Kikkawa & Anderson 1986, Kaitala *et al*. 1996).

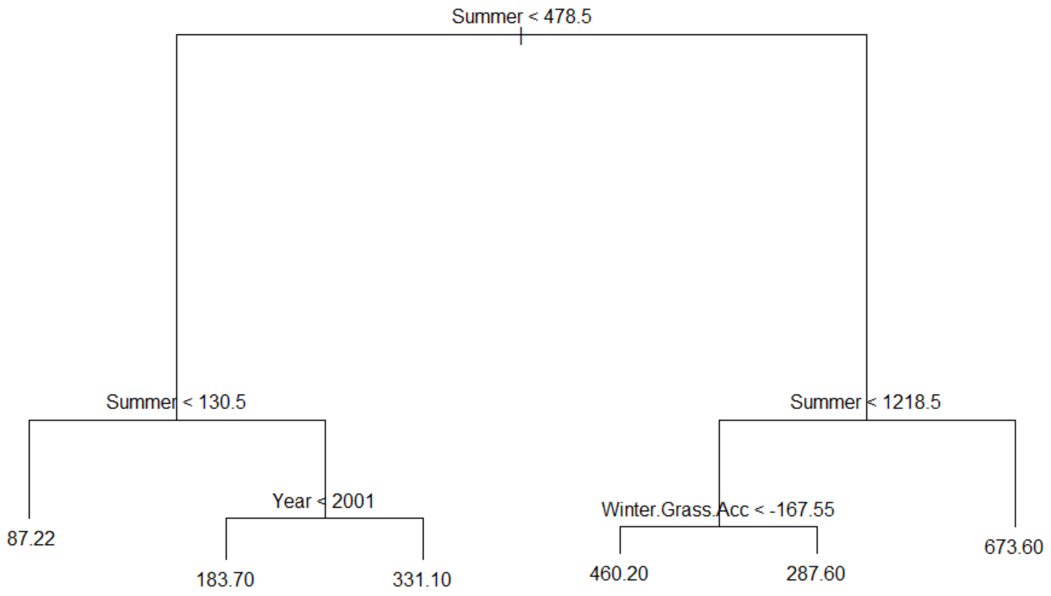


Fig. 2.27: Regression tree analysis of spring counts of *Brassicogethes aeneus* (beetles caught before 1st June) at Rothamsted and Broom's Barn. Each of the five branching points (nonterminal nodes) is labelled with the explanatory variable and value of that variable that determines the split. For each of the six terminal nodes, the mean of the observed values of Spring *B. aeneus* counts is given.

2.6.2 Pattern of increase within year

It appears that the increasing trend in *B. aeneus* abundance is found across the year with only early season (March/April Fig. 2.14) and July (Fig. 2.17) trends differing from the overall trend (though it appears that August/September counts are increasing at a greater rate (Fig. 2.18)). The different pattern in July is hard to explain. It could perhaps be a function of delayed or advanced emergence of the Newly Emerged Generation or related to the timing of oilseed rape harvest in different years.

2.6.3 Context of pollen beetle increase against other brassica-feeding pests

The increasing overall trend in *B. aeneus* is in alignment with the cabbage stem flea beetle (*Psylliodes chrysocephala* L.) which has shown large increases in larval abundance in recent years, although they may have peaked in 2015 (Collins 2017). Conversely, these beetle trends are contrary to the trend in brassica-feeding butterflies in England, with the small white butterfly (*Pieris rapae* (L.)) and large white butterfly (*Pieris brassicae* (L.)) both showing non-significant declines between 1990 and 2018 (DEFRA 2019b). Overall, these different findings demonstrate that the complex interactions within an ecosystem mean that there is large variability between species in magnitude and/or direction of effects of any individual driver of change (e.g. Tylianakis *et al*. 2008).

2.6.4 Conclusions

The relationship between spring and summer generations – both within and between years is a promising finding, indicating that there may be potential to forecast numbers migrating in the Spring - the period when crops are most at risk – based on counts taken in the summer of the previous year. Given that there is a large amount of scatter in this relationship any conclusions drawn must be viewed as tentative, especially as the GAMs indicate that the meteorological parameters, including air temperature during March-May are important, if not significant, variables (Table 2.6). If this finding does prove sound this would mean that counts taken from suction-trap samples by August could be used to make an initial assessment of the size of the *B. aeneus* migration the following spring, providing an early signal to growers as to the potential for crop infestation and therefore allowing growers to plan and adjust their management strategies in advance.

Chapter 3.

Investigation of the daily abundance of pollen beetles in suction-traps

Chapter Summary

Daily data of Brassicogethes aeneus *pollen beetles from the Rothamsted suction-trap network 1987-2018 were investigated using four modelling techniques in order to determine the best method for modelling daily abundance.*

It was found that generalized additive models were unable to adequately model the dataset due to the very large number of zero counts. Zero-inflated generalized linear models and two machine learning techniques; random forest and artificial neural networks did produce satisfactory models although none were able to forecast the sporadic high peaks in the counts that occur in the data.

Drivers of daily abundance were identified, with time of year, temperature, wind speed, solar radiation and the numbers of B. aeneus *recorded in the suction-traps after June 1st the previous year all considered important by one or more models. These results indicate that an accurate prediction of migration events that would be necessary for an improved decision support system for oilseed rape growers may not be possible. However, the conditions that favour migration events and lead to an increased likelihood of such can be determined.*

3.1 Introduction

3.1.1 Pollen beetles as pests

Pollen beetles (*Brassicogethes aeneus* (Fab.)) are the most commonly found insect in oilseed rape (OSR) crops in Europe (Williams & Free 1978, Valantin-Morison *et al*. 2007, Williams 2010), where they can be a pernicious pest leading to significant yield loss (Finch *et al*. 1990, Williams 2010). *Brassicogethes aeneus* migrates to oilseed rape crops in spring to feed and oviposit. They chew into the unopened buds to feed on the pollen within, causing damage which leads to bud abscission and consequent seed yield loss (Free & Williams 1978, Winfield 1992). As winter OSR is beyond the susceptible stage by the time the majority of beetles emerge from hibernation impact usually is minimal compared to spring OSR (Williams 2010). However, on the continent mass migrations have caused significant yield loss (Zlof 2008). In spring OSR, however, there can be significant yield loss to *B. aeneus* as the beetles arrive in the crop at an earlier, more susceptible, stage of flowering (Finch *et al*. 1990, Alford *et al*. 2003, Hansen 2004).

3.1.2 Rise of insecticide resistance

Control of *B. aeneus* and other pests of OSR in Europe has been predominantly via synthetic insecticide sprays, often in response to a threshold set by the relevant regulatory body (Thieme *et al*. 2010b). Resistance to synthetic insecticides in *B. aeneus* was first reported from Poland in 1965 (IRAC 2006, Thieme *et al*. 2010b), with significant correlation recorded between the proportion of OSR in the landscape and resistance to pyrethroid insecticide (Riggi *et al*. 2016), the number of spray applications or the year of development of resistance (Richardson 2008). Resistance in *B. aeneus* to a range of insecticides, including all classes of pyrethroid, is now widespread across

Europe (Hansen 2003, Slater *et al*. 2011, Zimmer & Nauen 2011a, Nauen *et al*. 2012, Węgorek *et al*. 2009, Palagacheva 2014, Zimmer *et al*. 2014, Heimbach & Brandes 2016) with a concomitant loss of effective control measures.

3.1.3 Reduction of prophylactic spraying

Growers need to move towards more sustainable systems and find approaches that can reduce the current level of crop inputs. Eliminating prophylactic spraying of crops is an important tactic to reduce the spread of resistant strains and facilitate a move away from an over-reliance on insecticides. This tactic requires an element of risk on behalf of the grower as not spraying when necessary may result in lower yields, whereas prophylactic actions are relatively cheap in costs and labour. To encourage growers to reduce prophylactic sprays a range of alternative methods have been, and continue to be, developed (Cook & Denholm 2008). Methods include trap crops (Hokkanen *et al*. 1986, Hokkanen 1989, Cook *et al*. 2004b) and other push-pull strategies (Cook *et al*. 2007, Balaso *et al.* 2019), deployment of semiochemicals (Mauchline *et al*. 2018) and developing host plant resistance (Hervé & Cortesero 2016) or reducing the attractiveness of the crop (Cook *et al*. 2013).

3.1.4 Decision support systems

An important, more sustainable, option in pest control is the use of decision support systems (DSSs) that assist the grower in determining whether weather conditions and crop growth stage necessitate crop inspection and/or chemical control. For *B. aeneus* the primary commercially available DSS in use until recently in Europe was ProPlant (www.proplant.com), now subsumed into the Xarvio DSS platform (www.xarvio.com) which used accumulated spring temperatures in combination with weather parameters (daily mean & maximum temperatures, wind speed and sunshine) to forecast *B. aeneus* phenology (i.e. the start, end and peaks of *B. aeneus* migration) (Johnen *et al*. 2010, Junk *et al*. 2016). Furthermore, the DSS helps to optimise the timing of monitoring to detect breaches in treatment thresholds. Ferguson *et al*. (2016) showed that this system can reduce monitoring effort required to detect thresholds by half. proPlant was recently acquired by BASF as part of the Xarvio agricultural decision support package (Ryan 2019) and it is currently unclear what access there will be in the future. Importantly, whilst current *B. aeneus* DSSs for growers can predict the timing of the spring arrival into crops 3-5 days in advance (Johnen *et al*. 2010, Ferguson *et al*. 2016), they cannot forecast the scale (abundance) of the immigration and therefore the likelihood that thresholds will be breached. Furthermore, there are no DSSs currently supported for spring OSR as the metrics used (start of migration, good migration conditions, percentage migration) do not provide the required information for the time period that spring OSR is vulnerable. If these aspects were available then it would be possible for growers to assess both when, and how important it is in any given year to monitor both winter and spring OSR crops for beetles.

3.1.5 The Rothamsted Insect Survey suction-trap network

The Rothamsted Insect Survey suction-trap network (Storkey *et al*. 2016) may be able to provide data that improves these forecasts. This network (described in detail in Chapter 2) was initially designed to monitor aphid migrations (Bell *et al*. 2015) and forecast models which include abundance as well as phenology have been developed (Harrington & Woiwod 2007, Bell *et al*. 2015, Fabre & Dedryver 2017). RIS suction-traps (Macaulay *et al*. 1988, Bell *et al*. 2015) sample aerial insect populations at 12.2m above ground level. Samples are collected daily, aphids removed, and the bycatch stored in 95% ethanol with 5% glycerol added (Shortall *et al*. 2009). The bycatch has been stored since 1974 (Shortall *et al*. 2009) and previous work on *Culicoides* biting midges in suction-traps (Sanders *et al*. 2011) resulted in the development of a seasonal abundance model allowing the prediction of daily *Culicoides* abundance according to time of year and weather conditions (White *et al*. 2017) thereby demonstrating the possibility of using suction-traps to develop this aspect of *B. aeneus* DSS tools. If it were possible to model *B. aeneus* seasonal abundance, the improved accuracy of forecasts beyond the mere phenological models currently available would be of great benefit to growers, reducing the need for monitoring in seasons where the abundance of beetles is likely to be low and therefore unlikely to breach spray thresholds.

3.2 Aims and objectives

The main aim of this work was to produce a farmer Decision Support System (DSS) model that captures the population dynamics and phenology of *B. aeneus* in the UK.

The objectives were;

1) to utilise suction-trap data to evaluate how different modelling approaches can best describe *B. aeneus* daily abundance both within-year across sites in England, and historically at two suction-trap sites.

2) to explore whether these models can:

    i) use meteorological data to build a predictive model of pollen beetle activity.

    ii) improve the range of decision-support forecasts in order to advance the timing of forecasts beyond the current three days.

3) to improve the knowledge of pollen beetle dynamics beyond the spring migration to provide information on risk to growers of spring OSR.

3.3 Data Collection

3.3.1 Pollen beetle abundance data

Data from the Rothamsted Insect Survey were used for this study. Samples from 12 English suction-traps were investigated (Fig. 3.1) gathering long-term datasets from two sites to allow investigation into long-term trends, and nine shorter datasets to provide a wider geographical spread:

- Rothamsted (1987-2018) – 32 years daily data counts
- Broom's Barn (1989-2018) – 30 years of daily data counts
- Nine sites with shorter datasets (2015-2018) – 9 x 4 years of daily data counts
- Preston, providing data only for 2015.

Samples are collected daily for the duration of the aphid season which typically runs from the end of March to the end of October, and weekly during the winter and early spring. All samples between the last week of February and the first week of September were investigated for this study. *Brassicogethes aeneus* beetles were identified and counted. Weekly samples from early in the year were divided by seven to produce daily figures. Note that *B. aeneus* is difficult to separate quickly from the closely related *B. viridescens*, however the latter is very rare in the samples and thus misidentification is not thought to have a significant impact on counts.

To explore the effect of the abundance of beetles entering hibernation, the Newly Emerged Beetles (See Chapter 2 Section 2.4.6) from the previous year were included as a potential explanatory variable.

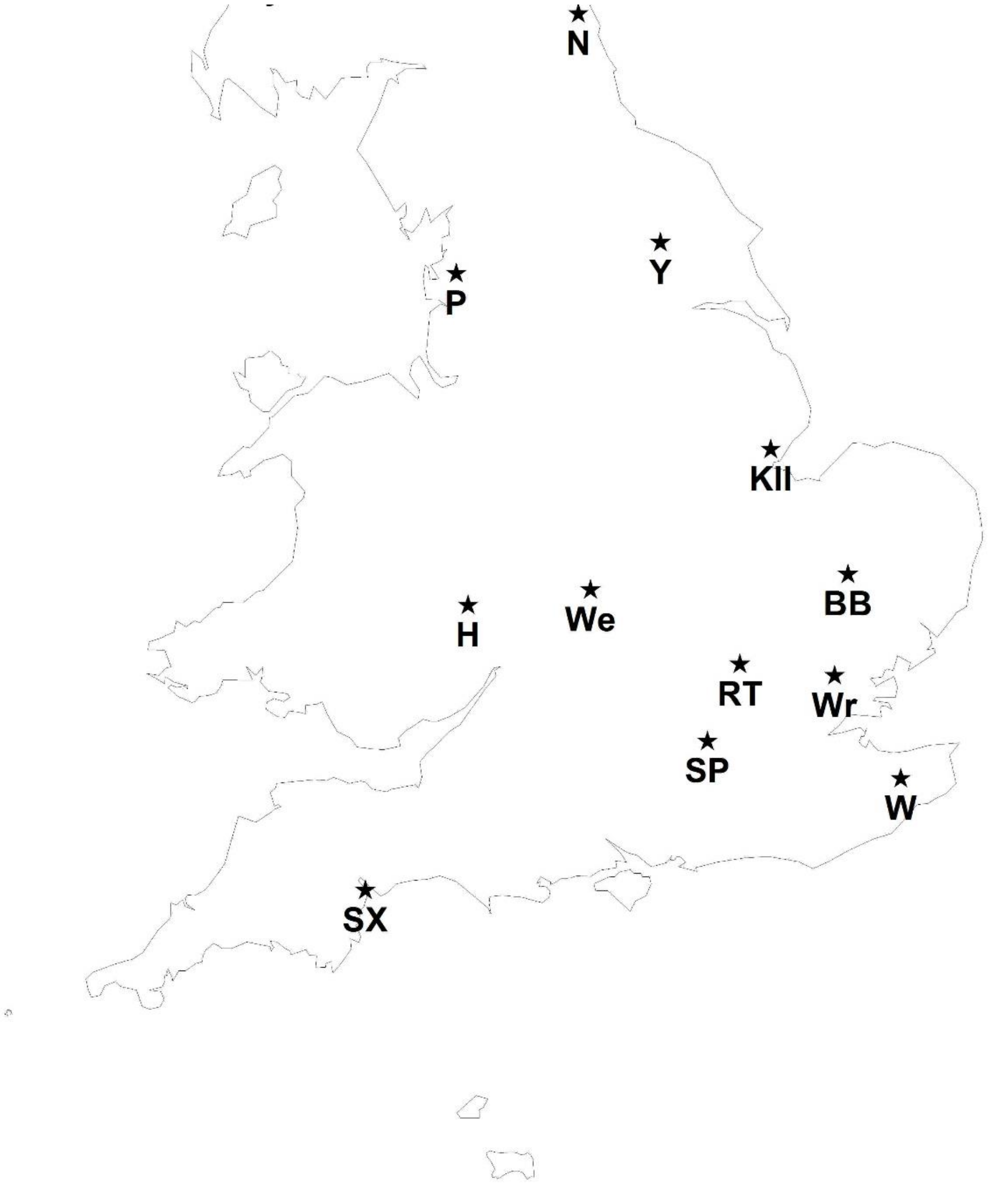


Fig. 3.1: The RIS suction-trap network in England. Site codes, from north to south are: N: Newcastle, Y: York, P: Preston, KII: Kirton II, BB: Broom's Barn, We: Wellesbourne, H: Hereford, RT: Rothamsted Tower, Wr: Writtle, SP: Silwood Park, W: Wye, SX: Starcross.

3.3.2 Meteorological data

Meteorological parameters that may have an effect on the daily abundance of *B. aeneus* were selected for investigation. Table 3.1 provides a summary of the selected variables. According to Dormann *et al*. (2013) a good strategy in variable selection is to collect ecologically relevant data that are feasible to collect and as close as possible to the mechanism (in the variable sequence: resource-direct-indirect-proxy, Austin 2002). Here, variables were selected as those that may predominantly have either an indirect effect on the beetle life history (i.e. temperature, dew point, humidity), a direct effect on their life history or behaviour (i.e. temperature, sunshine, precipitation, wind speed) or affect trap efficiency (e.g. wind speed). There is evidence that barometric pressure has an effect on insect behaviour (Fournier *et al.* 2005, Pellegrino *et al.* 2013, Musiolek & Kočárek 2016, Miao *et al.* 2021, but see Tansey *et al.* 2010). However, most research points to a change in pressure being the important cue, particularly for flight activation (Fournier *et al.* 2005, Pellegrino *et al.* 2013, Austin *et al.* 2014, Martini & Stelinski 2017, Martini *et al.* 2018). As change in pressure is a metric impossible to capture at the scale used here it is discounted from these models.

3.3.2.1 Daily data

Meteorological data were sourced from NASA via their online portal at https://power.larc.nasa.gov/data-access-viewer/ as solar radiation (sunshine) was considered to be a potentially important parameter and historical data for solar radiation was not available from other data providers such as the UK Met Office at the required spatial granularity.

3.3.2.2 Winter weather

To capture the effects of winter climate on daily counts the following meteorological parameters were selected: *Winter Rainfall*, designated as the mean rainfall during the months December, January and February; *Winter Grass Temperature*, designated as the accumulated day degrees of the grass minimum temperature below 0°C during the months December, January and February; and *Autumn Rainfall*, designated as the mean rainfall during the months September, October and November. The rainfall parameters were selected to explore the effects of precipitation during overwintering (*Winter Rainfall*); and waterlogging of the soil prior to, or during, the beetles selecting their overwintering sites (*Autumn Rainfall*). Accumulated grass minimum temperature was selected to explore the effects of low temperatures at ground level where beetles are overwintering (*Winter Temperature*). These parameters, alongside spring temperature, have been shown (Chapter 2) to have an effect on annual counts at Rothamsted and Broom's Barn. The spring temperature from Chapter 2 was not used here as it is coincident with the daily temperature measures described above.

3.4 Analysis and Results

3.4.1 The challenge of suction-trap data

The *B. aeneus* daily suction-trap data are characterised by sporadic peaks of very high counts within a wider context of low to zero counts and often show day-to-day changes of two orders of magnitude (Fig. 3.2). This, combined with a lack of spatial synchrony (explored in Chapter 4), makes modelling the data using traditional statistical techniques a challenge. Here, I outline the steps taken to explore and analyse the data. All analyses were carried out in R versions 3.5.0 and 3.6.1 (R Core Team 2018) using RStudio Version 1.1.453 (Rstudio Team 2015).

Parameter Type	Parameter Name	Units	Derivation
Pollen beetle response	Number of *B. aeneus* per day	Count data	See Methods Chapter 2
Meteorological drivers	Minimum Temperature Maximum Temperature Mean Temperature Winter Grass Temperature	Degrees Centigrade	NASA
	Sunshine (daily average solar radiation)	MJ/m^2/day	NASA
	Precipitation Winter Rainfall Autumn Rainfall	Millimetres	NASA
	Relative Humidity	Percentage	NASA
	Dew point (daily average at 2m)	Degrees Centigrade	NASA
	Wind Speed (daily average at 10m)	m/s	NASA
Geographical drivers	Location	Degrees Latitude/Longitude	RIS data
Pollen Beetle driver	Number of *B. aeneus* in previous season	Count	See Methods

Table 3.1: Data and data sources used for models.

3.4.2 Data exploration

Graphical representation of the daily data for all sites are presented in Appendix B, for illustration Fig. 3.2 shows the daily counts at Rothamsted from 1987 to 2017. The annual variability is clearly high (see also Chapter 2, which shows that counts at Rothamsted started quite high, dropped to low values through the 1990s before rising to the highest counts in the late 2010s). In addition, within-year variability is also high, with spikes of high counts, possibly indicating single day migration events, found among generally low numbers.

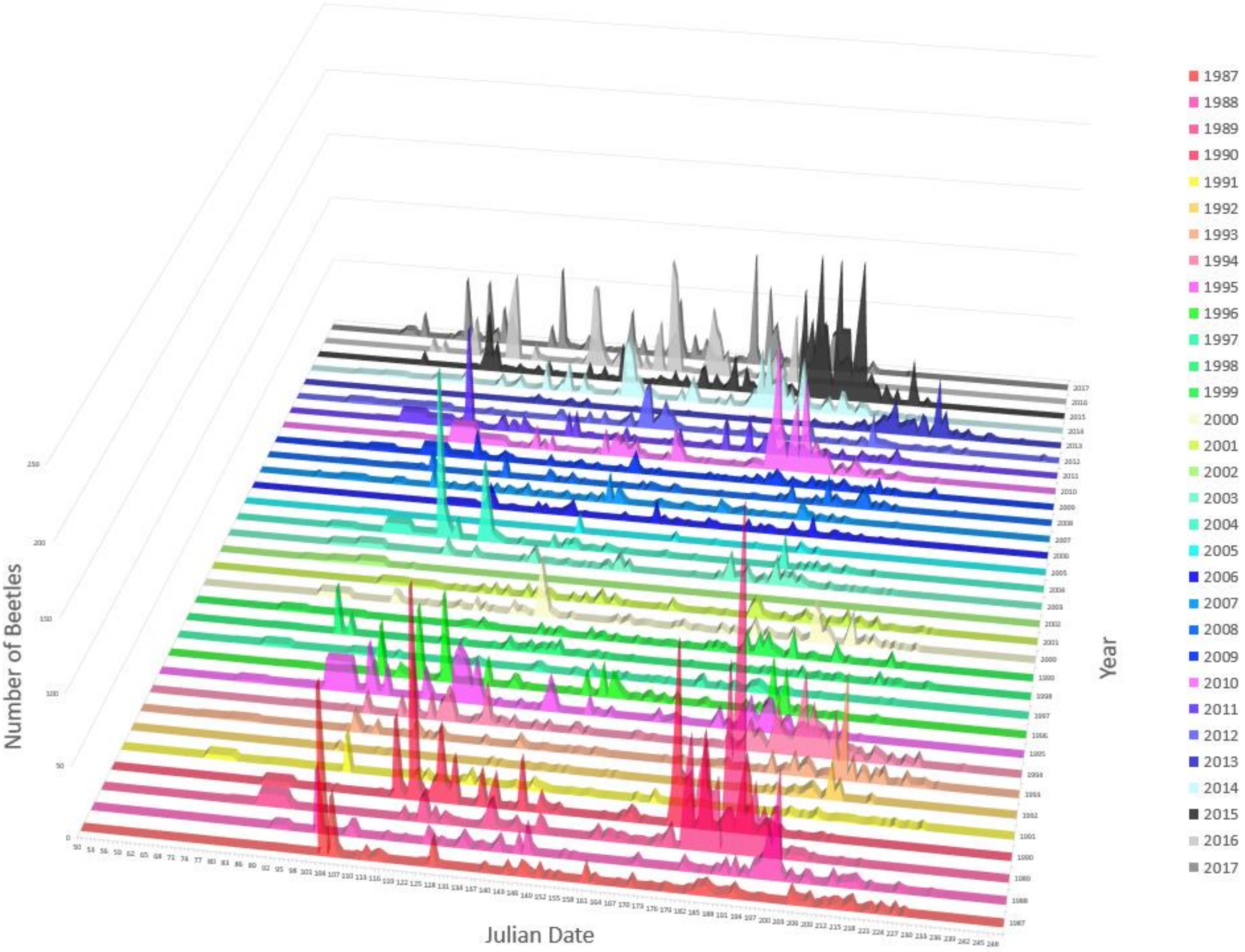


Fig. 3.2: Daily *Brassicogethes aeneus* counts – Rothamsted suction-trap 1987 to 2017.

3.4.2.1 Data distribution

In order to assess which modelling approaches would be appropriate to use the fitdistrplus package (Delignette-Muller & Dutang 2015) was used to establish the statistical distribution of the daily data. Due to the sheer size of the dataset only a subset of one site year was possible to test at one time. The best fit distribution for the data was a negative binomial distribution (Fig. 3.3), although this does not properly account for zero-inflation.

3.4.2.2 Principal Components Analysis

The relationship between variables was examined using Principal Component Analysis (PCA). Appendix A, Chapter 3 details how this was done. PCA is a method that can be used to simplify data by identifying the principal components within the matrix which then allows the removal of those variables that show likely collinearity, although this approach is not a formal test of it (*cf.* variance inflation factors). In short, a PCA views the variables in multi-dimensional space, with the number of dimensions being the number of variables in the dataset, and recasts within that multi-dimensional space to find the direction(s) that explain the most variability in the data. These new directions (known as eigenvectors) are assigned an eigenvalue based on the amount of variance in the data in that eigenvector. The eigenvector with the highest eigenvalue is therefore the principal component. Eigenvectors can then be ranked in a scree plot (Fig. 3.4) which allows consideration of the removal of components with low information (eigenvalue), and PCA biplots (Fig. 3.5) that show potential collinearity between variables when they lie along the same axis. Collinearity in predictors can be undesirable in least-squares regression due to the non-orthogonality (correlation) between them (Dormann *et al*. 2013). These problems are exacerbated in stepwise selection methods because if one, rather than another collinear variable is dropped the selection process is compromised and may proceed along the wrong path (Meloun *et al*. 2002, Dormann *et al*. 2013, Harrell Jr. 2015). However, it is important to note that there are some situations in which collinearity has limited effects within a modelling framework. For example, if a model is used to predict new cases within the range of sampled data this can be done reliably as long as the collinearity between variables remains constant (Dormann *et al*. 2013, Harrell Jr. 2015). On the other hand, extrapolation beyond the (e.g.) environmental or geographical range of a dataset could

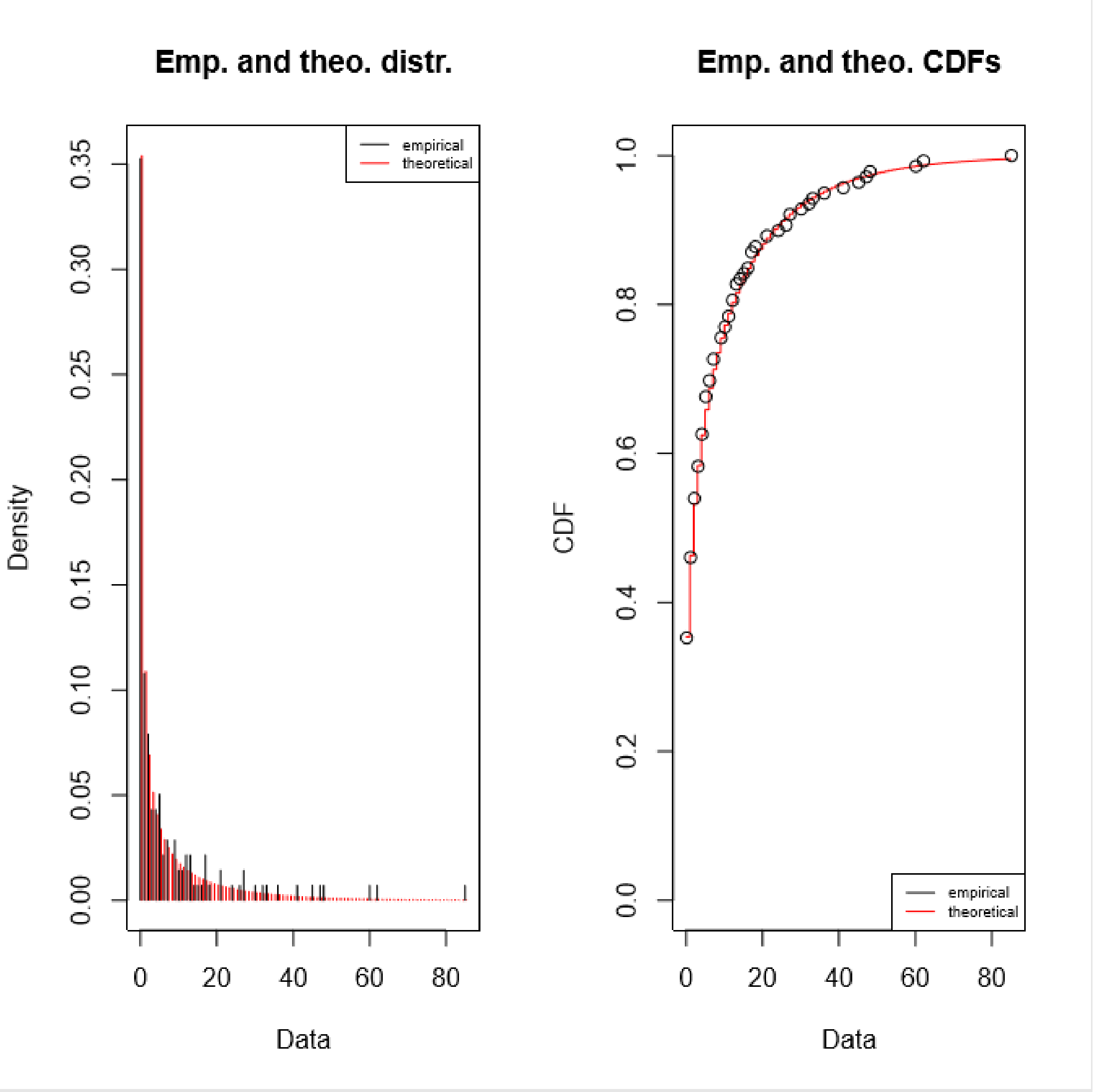


Fig. 3.3: Goodness-of-fit plot for daily data with a negative binomial distribution for 2017 *Brassicogethes aeneus* count data at Rothamsted. The observed distribution matches the theoretical distribution as expected by a negative binomial model.

be prone to serious errors as patterns of collinearity may not remain constant outside the sampled range (Dormann *et al*. 2013). PCA was carried out using the *tidyr* package in R (Wickham & Henry 2020), with graphics produced using *ggfortify* (Tang *et al*. 2016, Horikoshi & Tang 2016) and *factoextra* (Kassambra & Mundt 2020).

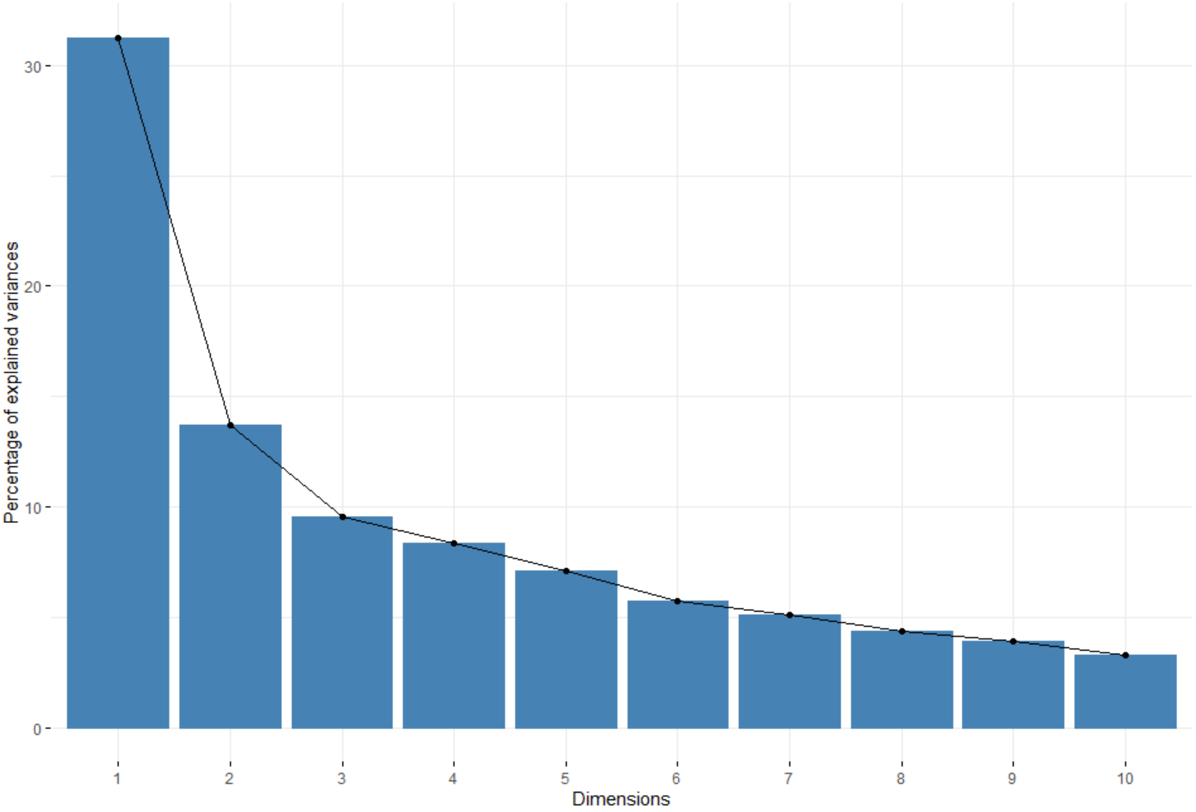


Fig. 3.4: Scree plot showing the percentage of explained variances (eigenvalues) of each dimension (principal component). Ideal scree plots are steep, then bend at an "elbow" which is the cut-off point between principal components that capture enough information to be useful and those that don't. Here the "elbow" is not as well-defined as an ideal plot, with the "elbow" situated across 2 and 4 principal components.

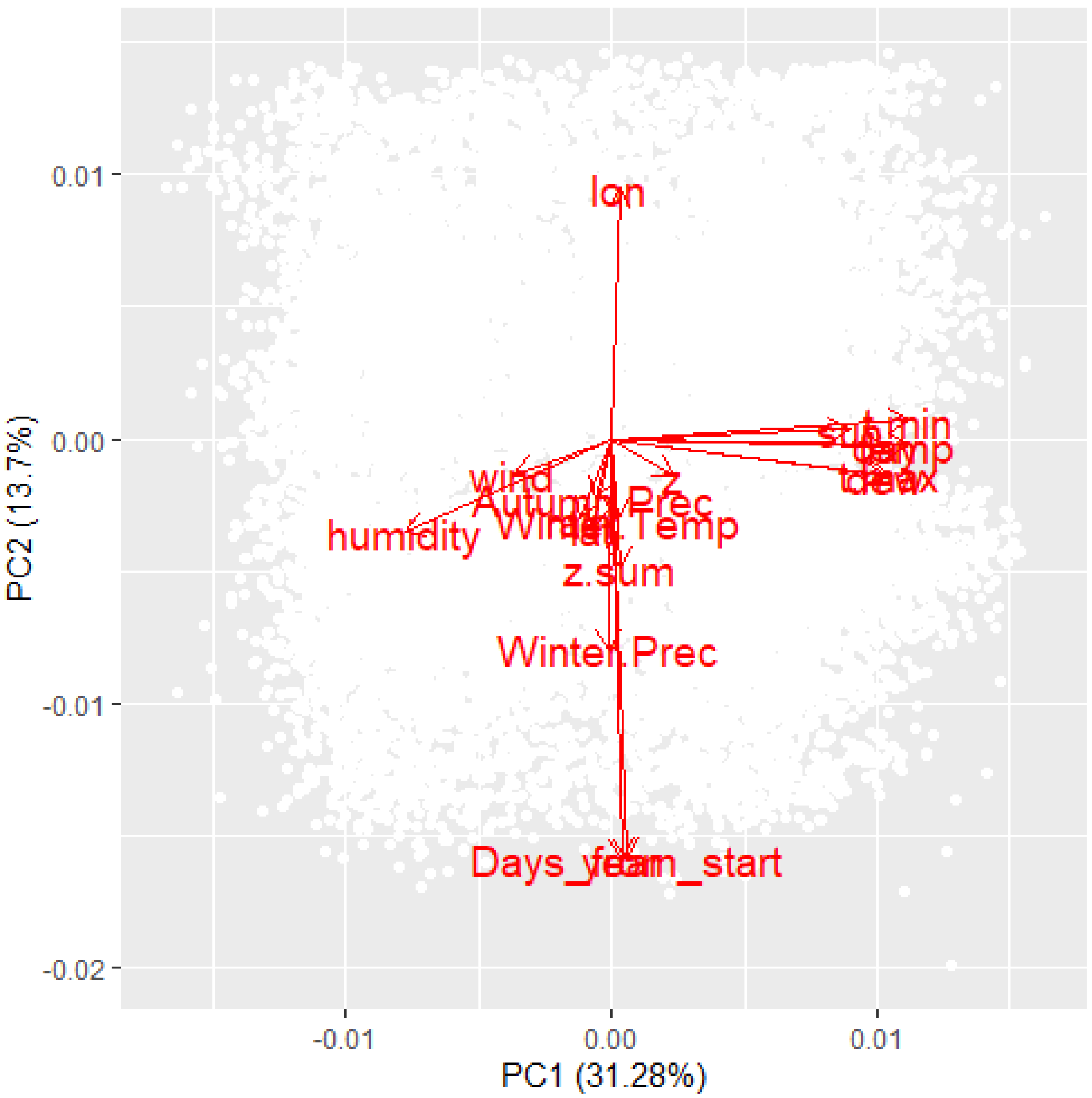


Fig. 3.5: PCA biplot of explanatory variables in the daily dataset. The cloud of are the location of each observation in the two dimensional space created by the first two principal components. The arrows show the strength (length) and direction (angle) of each variable's influence on the principal component. The longer the arrow the stronger the effect on variance of that principal component a variable has. Variables that are collinear are aligned in the plot, for example year and Days from start, humidity and wind and t.max and dew.

3.4.2.3 Excess zero issue

The dataset is characterised by a superabundance of zero counts, indeed even when all dates are removed post September 7th (the latest record in the dataset) the majority of daily counts are zero (Fig. 3.6). Unfortunately, even negative binomial regressions violate the assumption of normally distributed standardized residuals when there is a high proportion of zero scores (Lewsey & Thomson 2004).

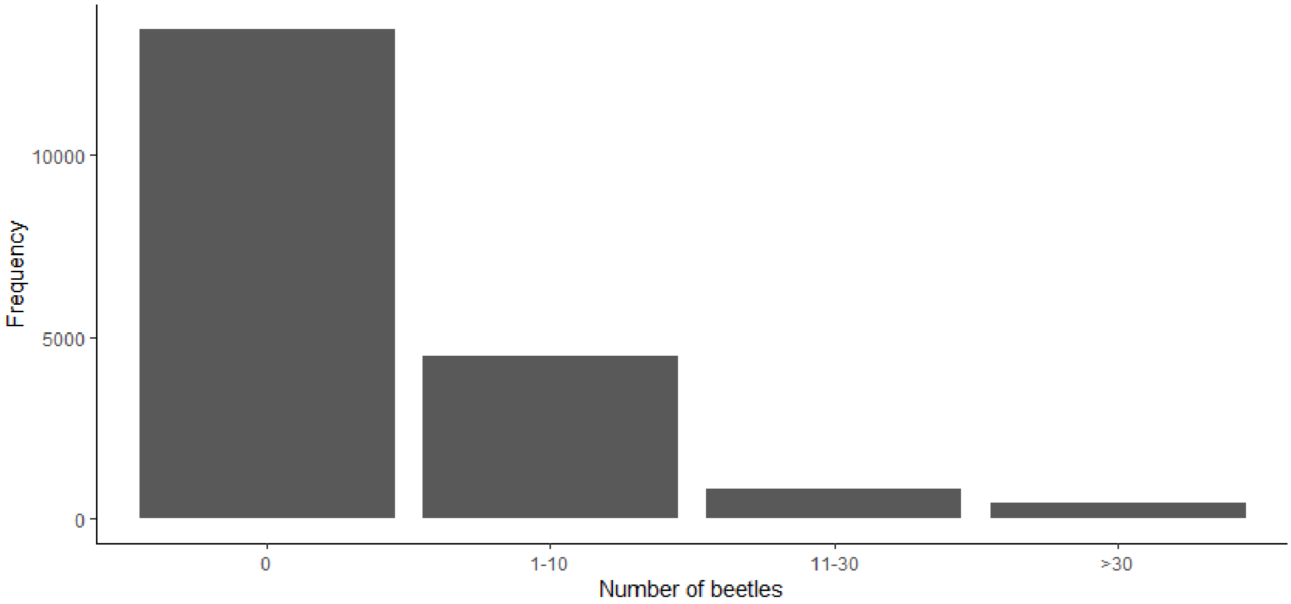


Fig. 3.6: Histogram of count classes in dataset of daily number of *Brassicogethes aeneus* caught in suction-traps. Most values are zeroes making it very difficult to model counts as a continuous variable.

3.4.2.4 Tree model

Given the high number of explanatory variables available and the possible collinearity issues identified in the PCA it was helpful to use a tree model to establish variable selection criteria *a priori.* Tree models are fitted using binary recursive partitioning to split the data and fit a simple prediction model within each partition. As a result, the partitioning can be represented graphically as a decision tree (Crawley 2007, Loh 2011, Song & Ying 2015). Each explanatory variable is assessed in turn with the variable explaining the greatest amount of deviance in the response selected. The deviance is then calculated to find the threshold in the explanatory variable which produces two mean values in the response – one above the threshold, the other below. This creates two "branches" to the decision tree which are then assessed in the same way with further partitioning where required until the most parsimonious tree is selected (Brieman *et al*. 1984, Ripley 1996, De'ath & Fabricius 2000). See Appendix A, Chapter 3 for detail on how this was accomplished in R.

The tree model (Fig. 3.7) suggested that, of the collinear temperature variables highlighted by the PCA, mean temperature is most important in explaining daily *B. aeneus* abundance. Day number (day) and sunshine (sun), which the PCA also showed as collinear with the temperature variables are also selected by the tree model as important variables. Other variables that the PCA highlighted as potentially problematical (wind/humidity and year/days from start) did not get selected as explanatory variables influencing *B. aeneus* catch by the tree model. These results should be accepted with caution, since multi-collinearity issues do not allow an absolute interpretation.

3.4.3 Generalized Additive Models

Generalized Additive Models (GAMs) were used to explore the relationship between daily *B. aeneus* abundance and a range of meteorological factors outlined above. GAMs are an extension of generalized linear models (GLMs) which fit smoothing terms with no *a-priori* selection of the parametric form of the function to be fitted (e.g. logarithmic, exponential, quadratic etc.). These non-parametric terms thus allow the shape of the relationship between response and drivers to be captured without the (possibly erroneous) presumption of a particular parametric form (Crawley 2007, Wood 2017).

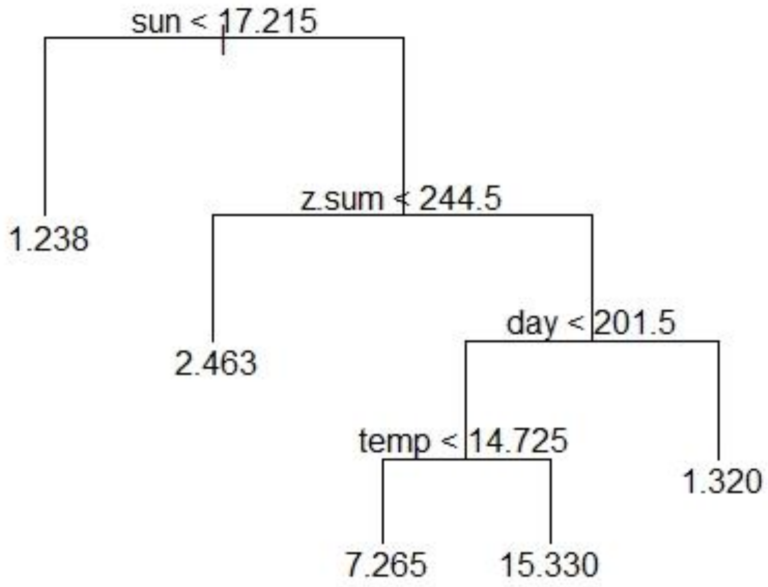


Fig. 3.7: Tree model for the explanatory values selected to explain the daily number of *Brassicogethes aeneus* caught in suction-traps. The mean number of beetles per day is given at the terminus of each branch, and the split in the selected variables is presented at the node. Here the model considers the amount of sunshine the variable that explains the greatest amount of deviance with the threshold being 17.2 MJ/m^2 of solar radiation below which the mean number of *B. aeneus* per day is 1.238, above this figure there are further splits based on z.sum (the number of beetles caught at the trap after July 1st the previous year), day number (with the threshold at day 201/202 – the 20th/21st July) and mean temperature.

3.4.3.1 Method used for GAMs

GAMs were assembled within the mgcv package (Wood 2017). After variable selection (section 3.3.2.2.5) the explanatory variables in Table 3.1 were selected, with humidity, rainfall and dew point removed, also latitude and longitude were combined as a spatial variable using a thin-plate spline (Wood 2003). All explanatory variables were smoothed using the restricted maximum-likelihood (REML) method that penalised overly-wiggly splines (Wood 2017). See Appendix A, Chapter 3 for more details on the model selected.

3.4.3.2 Variable selection

The final selection of explanatory variables based on the PCA and tree model exploration of the data was as follows: year was included and linked with latitude and longitude as a joint model term as a spatio-temporal variable. Day number was deemed important by the tree model and allows for within-year seasonal effects to be captured. Wind speed, a nuisance variable, was selected as it is known to have an effect on suction-trap efficiency (Taylor 1962) and *B. aeneus* flight activity (Skellern *et al*. 2017), its collinear variable as determined by PCA – humidity – was dropped. Accumulated winter temperature below 0°C, along with average autumn rainfall and average winter rainfall (with autumn defined as September, October & November and winter as December, January & February) were used to capture the effects of winter climate on overwintering beetle numbers, as in Chapter 2. Finally, the total count of beetles caught in the trap in the previous late summer (the Newly Emerged Beetles caught post-1st July) was included to represent the size of the previous year's population.

3.4.3.3 GAM results

Unfortunately, despite extensive efforts to select and tune the parameters, it proved impossible to create a GAM that did not either violate the statistical assumptions of a negative binomial distribution (Fig. 3.8) or reach the k-index values required for a complete model. This holds true whether the explanatory values are unmodified or scaled, indicating that it is the excess zeroes in the dataset that is the cause for the lack of model robustness.

3.4.4 Zero-inflated Generalized Linear Models

3.4.4.1 Introduction to Zero-inflated Generalized Linear Models

Excess zeroes are a common feature of ecological data (Barry & Welsh 2002). When a dataset contains an excessive number of zeroes, such that they exceed that allowed under the standard parametric family of statistical distributions it can be considered to have zero inflation (Tu & Liu 2014). Zero-inflated data need to be modelled carefully as the presence of excess zeroes can invalidate the underlying assumptions of the distribution within the analysis (Tu & Liu 2014, Zuur & Ieno 2016).

Zero-inflated generalized linear models (GLMs) are an extension of the frequentist GLM approach that models overdispersion and zero-inflation where the data are modelled in two steps (Barry & Welsh 2002, Zuur & Ieno 2016). First, the presence-absence component of the data is modelled via a GLM, with the logistic link recommended (Barry & Welsh 2002, Zuur & Ieno 2016). Second, the observed abundance, conditional on a value of the response greater than zero is modelled with either a truncated Poisson or truncated negative-binomial model (Welsh *et al*. 1996, Barry & Welsh 2002, Zuur & Ieno 2016).

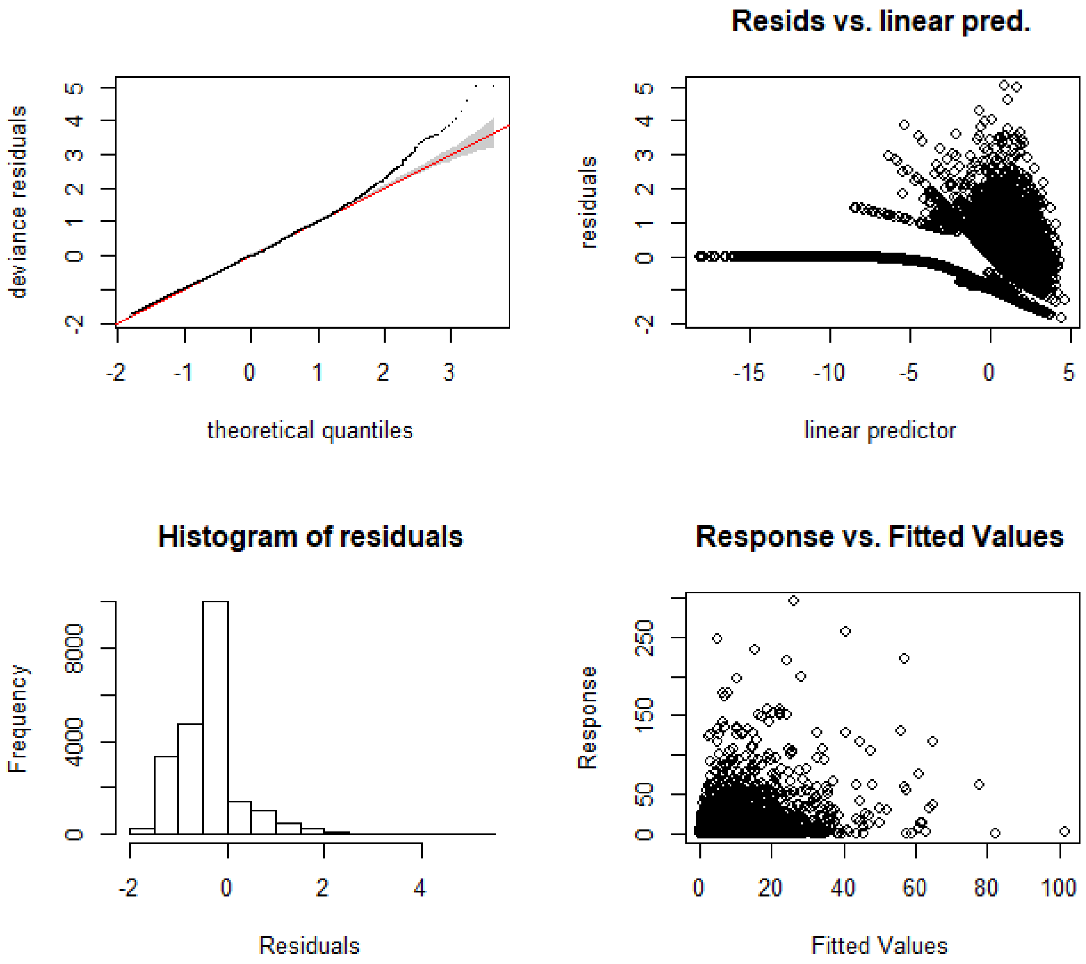


Fig. 3.8: Residual plots from the Generalized Additive Model (GAM) of daily *Brassicogethes aeneus* counts. These plots indicate that the statistical assumptions underlying the GAM are violated. The deviation in the Q-Q plot (top left) indicates that the right tail of the distribution in the GAM have deviated from the expected distribution in that they are more overdispersed than the model predicts, although otherwise well-behaved. The histogram (bottom left) is negatively skewed indicating that low counts are more frequent than the GAM predicts. The two biplots (right-side) indicate dysfunctionality in that the fitted values vs the response should be broadly positive and linear and there should be no pattern in the residuals v linear predictor, yet where observations were zero, a line intercepting 0 descending to 5 on the linear predictor scale, is clearly an artefact of the zero-inflation in the dataset.

3.4.4.2 Method for Zero-inflated GLM

The R packages *pcsl* (Zeileis *et al*. 2008b, Jackman 2020) and *MASS* (Venables & Ripley 2002) were used to construct the model, with *ggplot2* (Wickham 2016) and *ggeffects* (Lüdecke 2018) used to visualise the output. Given the wide range of values for each explanatory variable all values were scaled prior to modelling using the *scale()* function in R. Parameter selection was carried out using forward selection based on the relative importance of each individual explanatory variable as defined by the AIC and p-values from a zero-inflated linear regression on each variable (Table 3.2). Dew point, minimum and maximum temperature are all closely correlated with the more tractable mean temperature, likewise humidity correlated with wind speed (Fig. 3.5). Both wind speed and mean temperature were selected as variables that can easily be understood in terms of migration behaviour, with the other terms dropped. Latitude, temperature and precipitation in the previous seasons were low-ranking in terms of AIC and were excluded, even though they might be considered meaningful ecologically; Year and Days were excluded to avoid restricting the parameter space and allow for forecasting as far before the event as meteorological forecasts allow. See Appendix A, Chapter 3 for the modelling process in R.

3.4.4.3 Results

All terms in the model were significant (Table 3.3), more zeroes were expected when solar radiation was low, day number and wind speed was high and at lower longitudes (i.e. further west), with higher wind speeds and longitude (i.e. the further east) having a negative effect on count size, and higher temperature and the higher counts in the previous year having a positive effect (Table 3.3,

Fig. 3.9). The east-west count result was surprising as it would be expected that *B. aeneus* counts would be higher in the eastern regions where more OSR is cultivated (Chapter 2, Table 2.3). This result should, however, be considered in the context of the binomial part of the model which predicts more zeroes further west. This may mean that non-zero counts of *B. aeneus* are less regular further west, and when they do occur in the suction-traps it is because migration conditions are very favourable, leading to high counts. Alternatively, given that predicted values are generally underestimated compared to observed values (Fig. 3.10), it may be that estimates are lower in the east, given the greater number, and amount of variability, of non-zero counts, than in the west.

The underestimation of counts by the model shown in Fig. 3.10 is of concern, indicating that whilst a zero-inflated GLM does provide some clues as to what drives migration events in *B. aeneus* it may not provide enough predictive power to confidently forecast occasions when counts will be high enough to trigger crop inspections in a DSS.

	Degrees of freedom	AIC	Negbin intercept	Negbin coef	P	Binomial intercept	Binomial coef	Exponent coef (b)	P
Count of pollen beetles in the previous summer	5	57596.76	0.78005	0.557669	<0.001	-4.0649	0.53488	1.70725	<0.05
Julian day number	5	60944.8	0.82829	2.03963	<0.001	-14.389	11.8823	144677	<0.001
Solar radiation (sunshine)	5	61320.18	0.95013	0.55755	<0.001	-2.2303	-3.8365	0.02157	<0.001
Mean temperature	5	62350.51	0.93667	0.55695	<0.001	-3.4029	-4.6019	0.01003	<0.001
Minimum temperature	5	63160.26	0.37302	1.54494	<0.001	-9.2669	4.27986	72.2302	<0.001
Humidity	5	63319.12	1.07334	-0.3253	<0.001	-2.597	3.28633	26.7447	<0.001
Dew point	5	63550.75	0.99031	0.46138	<0.001	-2.4018	-2.816	0.05985	<0.001
Maximum temperature	5	63762.25	1.00123	0.41785	<0.001	-2.7917	-3.0046	0.04956	<0.001
Wind Speed	5	64548.55	1.11162	-0.3751	<0.001	-1.1467	1.23017	3.42183	<0.001
Days since Jan 1st 1984	5	65589.77	0.96633	0.2606	<0.001	-27.73	26.2626	2.5E+11	<0.001
Year	5	65668.11	0.95269	0.24345	<0.001	-50.296	47.6248	4.8E+20	>0.1
Longitude	5	65729.87	0.92082	0.28912	<0.001	-3.6406	-0.6084	0.54421	<0.001
Rainfall	5	65796.72	1.18012	-0.0299	<0.1	-1.3043	0.48606	1.6259	<0.001
Winter temperature preceding count	5	65811.29	0.92488	0.24297	<0.001	-4.3938	1.2596	3.52401	<0.1
Latitude	5	65868.35	1.00897	0.1227	<0.001	-2.9107	0.74858	2.11399	<0.001
Autumn precipitation preceding count	5	65872.31	0.95326	-0.2059	<0.001	-3.7577	-0.8686	0.41952	<0.01
Winter precipitation preceding count	5	65916.05	0.95289	0.00251	>0.1	-4.7834	1.1425	3.13459	<0.001

Table 3.2: Results of zero-inflated Generalized Linear Models of each explanatory variable, ranked by the Akaike Information Criterion (AIC) of each model. Negbin intercept and Negbin coef are the intercept and coefficient of the negative binomial portion of the model. Binomial intercept and Binomial coef are the equivalent terms for the binomial portion of the zero-inflated model, with the exponent given in Exponent coef (b) as exponentiating these coefficients place parameters on the odds ratio scale that estimates odds of not having beetles (zeros in other words).

Count model coefficients (negbin with log link):				
	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	0.78571	0.03211	24.47	< 2e-16 ***
Beetle count in previous year	0.52684	0.02225	23.68	< 2e-16 ***
Temperature	0.70084	0.03582	19.57	< 2e-16 ***
Wind speed	-0.52678	0.02951	-17.85	< 2e-16 ***
Longitude	-0.08125	0.025	-3.25	0.00115 **
Log(theta)	-1.37363	0.03428	-40.07	< 2e-16 ***
Zero-inflation model coefficients (binomial with logit link):				
	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-0.40973	0.08904	-4.602	4.19e-06 ***
Solar radiation (sunshine)	-2.13522	0.10518	-20.3	< 2e-16 ***
Wind speed	0.23727	0.04598	5.16	2.47e-07 ***
Julian day number	0.67251	0.05206	12.918	< 2e-16 ***
Longitude	-0.21479	0.04227	-5.081	3.76e-07 ***

Table 3.3: Summary statistics of zero-inflated Generalized Linear Model by model partition.

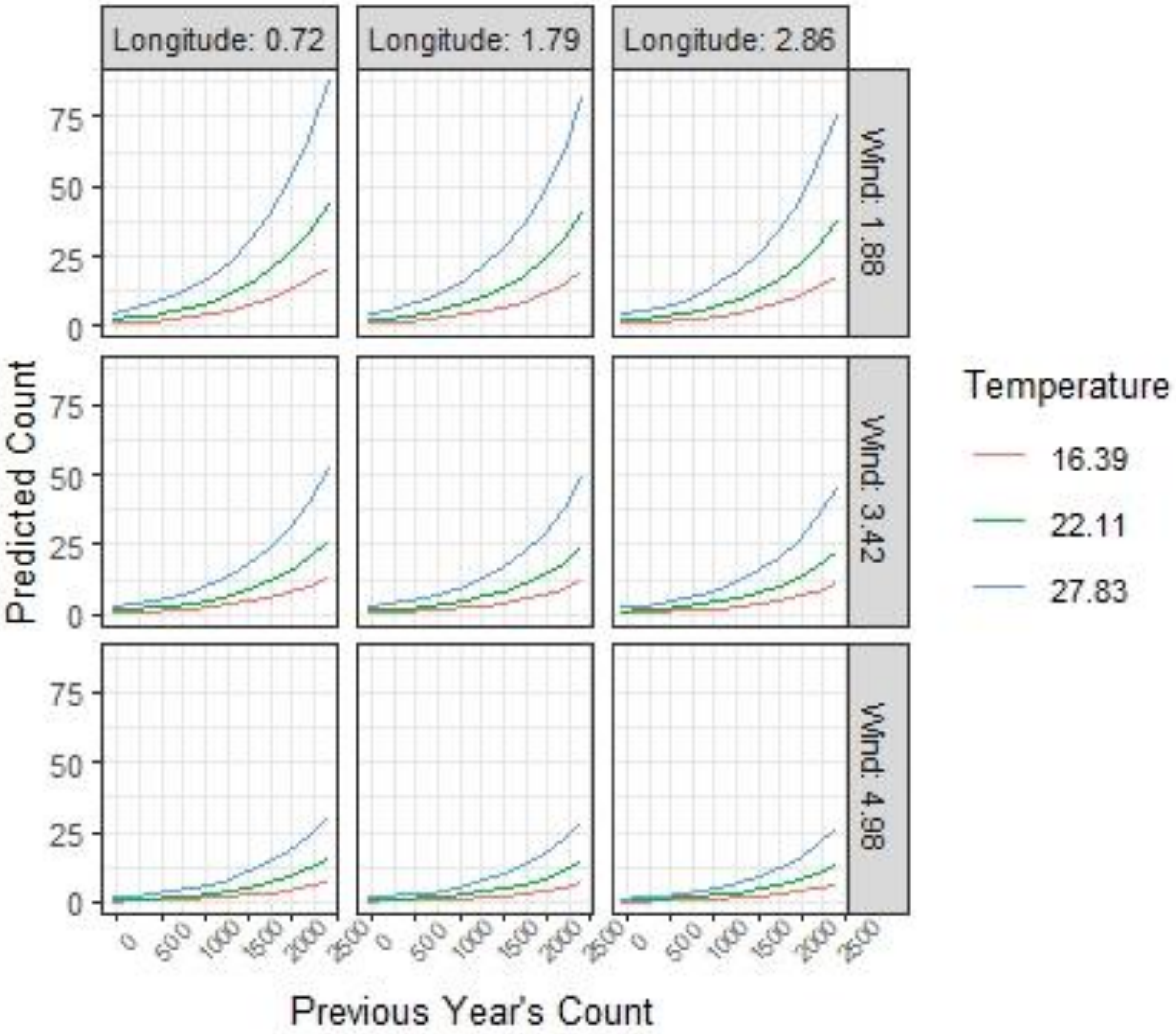


Fig. 3.9: Predicted non-zero counts of *Brassicogethes aeneus* in suction-traps given the count of beetles from June 1st the previous year at three levels of temperature (colours), three levels of wind speed (top to bottom) and three lines of longitude. Note, these values are scaled, not absolute. Predicted counts rise more or less exponentially with increasing numbers in the previous year; these counts are greater at higher temperatures, lower wind speeds and further west.

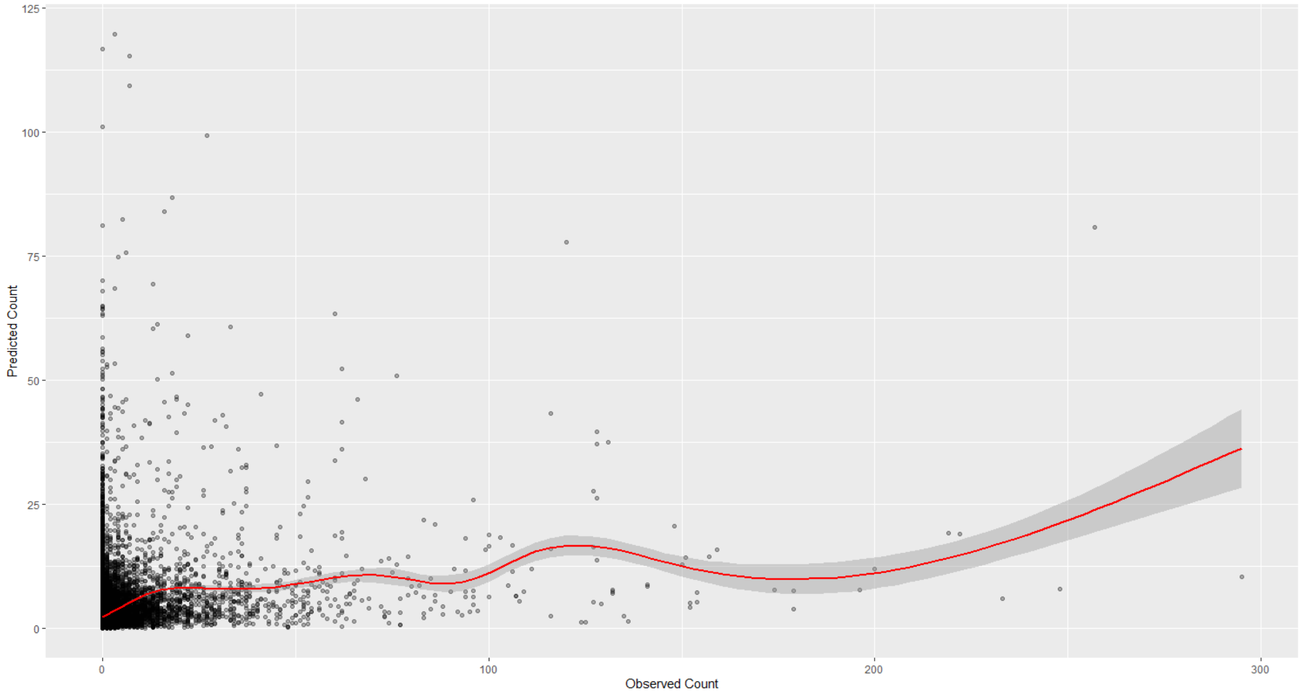


Fig. 3.10: Plot of observed versus predicted values from the zero-inflated Generalized Linear Model. Points denote individual counts, with the red line indicating the best fit spline with 95% error in grey. Counts are generally underestimated in the model compared to observations although there are several occasions where high (>100) counts are predicted when low (<50) numbers were observed.

3.4.5 Machine learning

Machine learning approaches are insightful and can be of use when more traditional statistical methods fail. Two approaches are investigated here: Random Forest, a supervised learning algorithm and artificial neural networks (ANNs) whose predictions arise from non-linearities and complex interactions of potentially numerous environmental covariates.

3.4.5.1 Random Forest

Random Forest is a machine learning method using a robust, iterative process to derive either a classification or regression trees (Breiman 2001, Cutler *et al*. 2007, Howard & Bickford 2014). When the response variable is continuous the resulting tree is termed a regression tree, with classification trees formed when the response data are categorical (Crawley 2007). A Random Forest analysis fits a series of trees to a dataset, combining the predictions from these to produce a predictive model based on ranked variables scored by predictive value (Cutler *et al*. 2007, Dinsdale *et al*. 2013, Coulthard *et al*. 2019). These models can be displayed in the form of a flow chart-style decision tree where successive nodes split the tree into branches based on the value of the parameter selected, much like the tree model presented above (Fig. 3.7). The explanatory variables can also be ranked in order of importance, for each explanatory variable the mean decrease in Gini coefficient can be calculated. The Gini coefficient is a measure of homogeneity that ranges from 0 (homogenous) to 1 (heterogenous). Each time an explanatory variable is used to split a node the Gini coefficient for the child nodes are calculated and compared to the original node, with variables showing a greater decrease in Gini coefficient deemed more important, in other words a “purer” split is favoured (Menze *et al*. 2009, Dinsdale *et al*. 2013). In ecology, Random Forest modelling has been used to investigate species interactions (e.g. Miller *et al*. 2014, Desjardins-Proulx *et al*. 2017, Pichler *et al*. 2020), species distribution modelling (e.g. Williams *et al*. 2009, Bradter *et al*. 2013, Shabani *et al*. 2016, Rowbottom *et al*. 2020), land-cover classification (e.g. Kane *et al*. 2014, Ahmed *et al*. 2015, Hao *et al*. 2015) and trait-based forecasting of species trends (e.g. Philibert *et al*. 2011, Holliday *et al*. 2012, Coulthard *et al*. 2019), however their use in time-series abundance modelling is less common (Shoemaker & Akçakaya 2014, Ward *et al*. 2014, Humphries *et al*. 2018).

3.4.5.1.1 Random Forest Method

The R statistical packages *RandomForest* (Liaw & Wiener 2015), *rpart* (Thernau & Atkinson 2018) and *partykit* (Zeileis *et al*. 2008a) were utilised. *RandomForest* was used to fit many classification trees to the dataset, combining the predictions from these trees to produce a decision tree with predictions based on variables ranked by predictive value (Cutler *et al*. 2007). Each model run generated 5000 trees with each run having an ‘mtry’ parameter (the number of variables available for splitting at each tree node) of four (the square root of the number of variables). There is some discussion as to the suitability of *RandomForest* for analysis that includes variables that vary in their scale of measurement as they do with the *B. aeneus* dataset (Strobl *et al*. 2007), however the tree partitioning algorithm should be independent of scaling for most circumstances, since they are merely attempting to establish thresholds within the data. Each model was run ten times (creating 50,000 trees) to obtain an estimate of variation in variable importance. The code for this model is shown in Appendix A, Chapter 3.

3.4.5.1.2 Random Forest results

Julian day number was deemed the most important of the selected variables (Fig. 3.11) with sunshine, temperature and wind speed also ranking highly. Latitude and longitude were deemed the least important, indicating that, for the Random Forest model the spatial element of the explanatory variables is secondary to the meteorological and temporal elements. It is also worth noting that for this model the parameters derived from previous weather or recording events (the count of beetles in the previous year (noted as z.sum in Fig. 3.11), autumn and winter precipitation and winter temperature) are ranked below the variables recorded on the day of capture (Fig. 3.11).

Decision trees created using the variables from the random forest revealed Julian day number, minimum temperature, wind speed, longitude, sunshine and the count of beetles after June 1st the previous year to be the most important factors in determining daily (Fig. 3.12). Zero counts were considered the most likely at eight of the eleven terminal nodes, with low (<10) counts

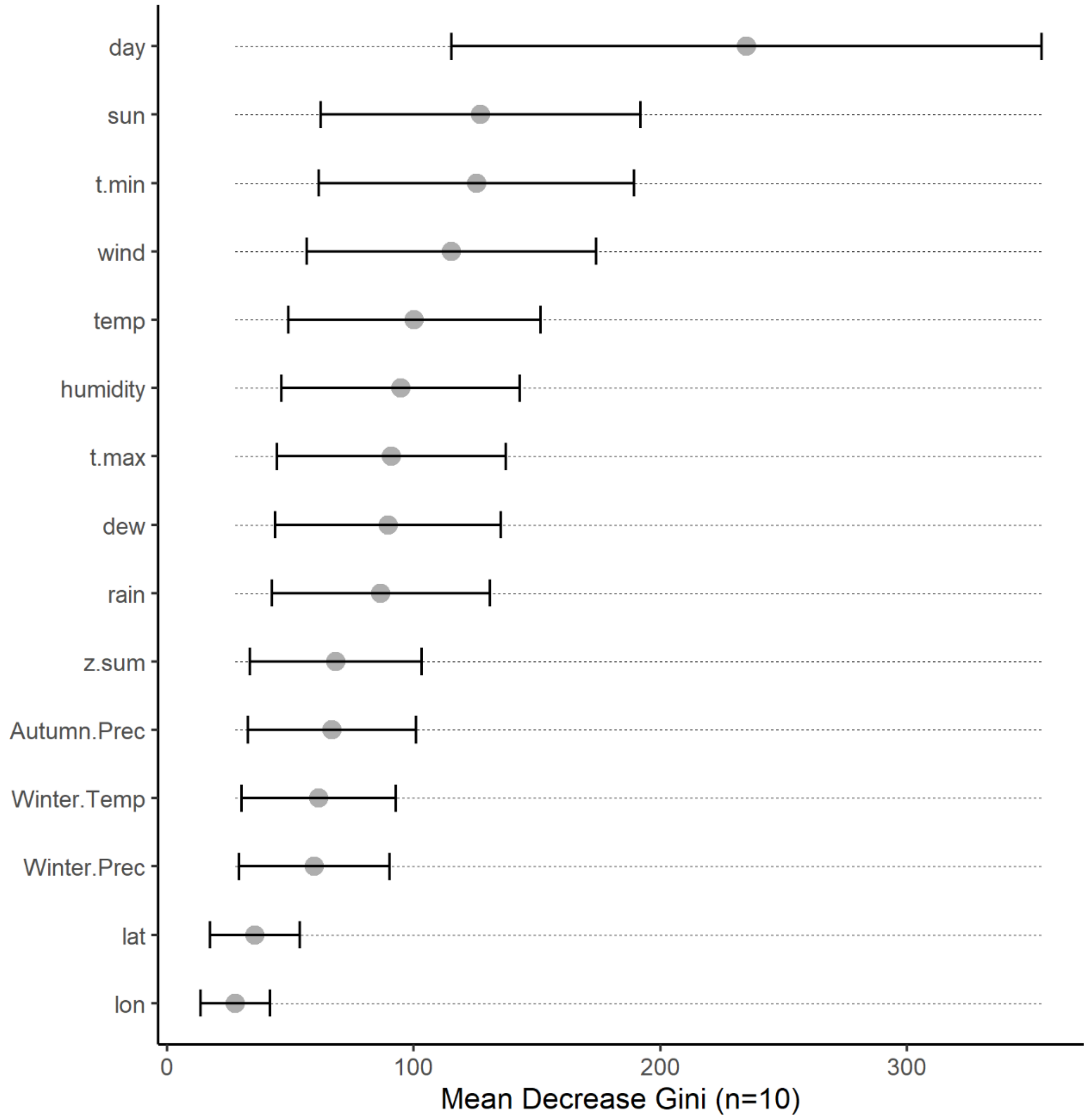


Fig. 3.11: Importance rank of the explanatory variables, using mean (+/- Standard Error) decrease in GINI index. Julian Day number ('day') is likely to be the most important variable, with latitude and longitude (lat, lon) the least important.

the most likely at the other three. Counts of less than ten were the second most likely outcome in all occasions when they were not the first (Fig. 3.12).

Day number is the first consideration in the random forest model (Fig. 3.12) with days before March 8th (Day 67, Node 1) and after July 29th (Day 210, Node 3) highly likely to produce zero counts (Nodes 2 & 4). Between these dates if the minimum temperature is below 11.965°C (Node 5) then, again, zero counts are likely, although there is a ~20% chance of a count between 1 and 10 (Node 6). If minimum temperature exceeds 11.965°C (Node 5) then wind speed is considered (Node 7), with values above 4.005 meaning that just over 60% of the time zeroes, and just over 20% of the time values of 1-10 are recorded (Node 8). If average wind speed is lower than 4.005 m/s (Node 7) and the date is between March 8th (Node 1) and April 8th (Day 98, Node 9) then it is more likely than not that a count of 1-10 will be recorded (Node 21). Between April 8th (Node 9) and July 29th (Node 3) at lower wind speeds the value for the daily count rests on the count of *B. aeneus* after July 1st the previous year (Node 10), with counts higher or lower than 244 having outcomes that depend on different parameters. If the previous year's count is lower than 244 (Node 10) then an average wind speed of greater than 2.585 m/s (Node 11) will lead to Node 12 which predicts either a zero or 1-10 count in most cases, if winds are lighter then sunshine becomes important, with less than 11.39 MJ/^{m2} (Node 13) leading to Node 14 where 0-10 *B. aeneus* are likely, and higher than 11.39 MJ/^{m2} (Node 13) meaning a positive count is more likely than a zero, though low counts predominate (Node 15). If counts in the previous year are higher than 244 (Node 10) then more westerly traps (Newcastle, York, Preston, Hereford, Wellesbourne and Starcross, Node 16) will fall in Node 17 where zero counts are most likely but counts of 10-30 or even higher have a larger chance of occurring than most other nodes. Traps east of longitude -0.667 (Kirton, Broom's Barn, Rothamsted, Writtle, Silwood Park and Wye, Node 16)) will terminate in either Node 19 or Node 20 depending on whether minimum temperature is above or below 14.355°C (Node 18), these nodes

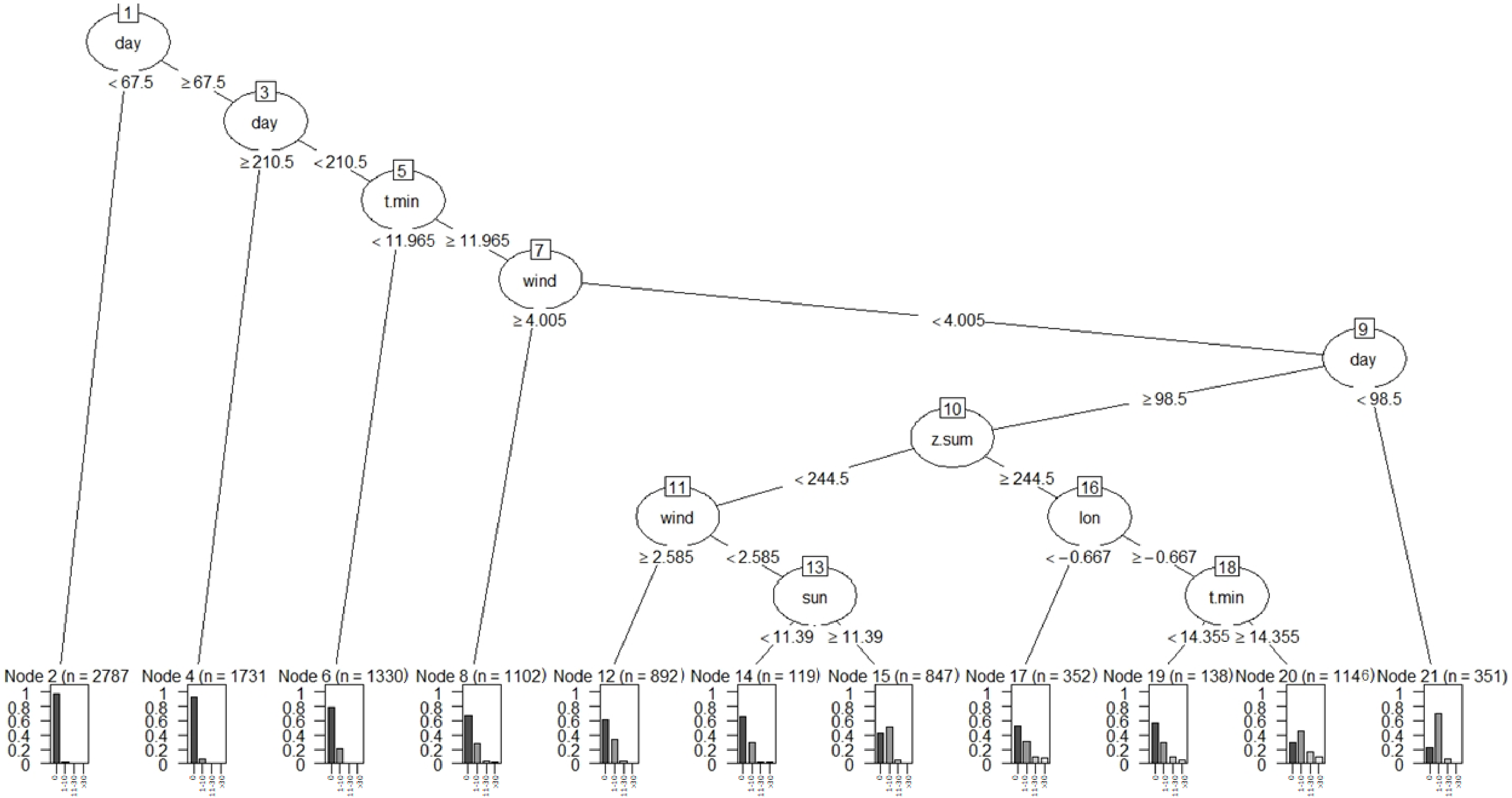


Fig. 3.12: Decision tree for the Random Forest model of daily suction-trap catches. The 21 nodes are divided between 'splits' where decisions that minimize the misclassification error in each leaf are made based on the thresholds of the explanatory variable under consideration (e.g. in the first split, the threshold is day 67.5 i.e. between March 8th and 9th) and 'leaves' where the breakdown of predicted counts is presented in a bar chart. Thus, if day is earlier than day 68 (March 9th) Node 2 applies where the bar chart indicates that nearly all predicted counts will be zero, with very few in the 1-10 category. In fact, zero will be predicted most of the time at all nodes except Nodes 15, 20 and 21, in which 1-10 is the most commonly predicted count.

return a complex pattern of predicted values with all count levels potentially returned at different levels of likeliness (Fig. 3.12).

3.4.6 Artificial Neural Networks

3.4.6.1 Introduction to Artificial Neural Networks

Artificial Neural Networks (ANNs) are composed of multiple processing layers which learn data representations with multiple layers of abstraction (LeCun *et al*. 2015). The allusion to 'neural networks' infers that computers perform like interconnected brain cells, finding patterns through decision-making. Just like the brain, in which neurons are clustered according to the task, ANNs have layers in which particular problems are solved and then passed on to other layers for completion. At its most basic, an ANN has three layers: input (training data), hidden (computation and processing) and output (results) and differs from standard statistical analyses in that there are few *a priori* assumptions. Instead ANNs promote self-adaptive and self-organising behaviours that use weights and biases, to capture subtle functional relationships within the data that are driven by unknown, or hard to determine processes, a compelling example of machine-learning to solve complex problems (Agrawal & Mehta 2007). Fig. 3.13 presents a schematic of a neural network in which the data are randomly split into two subsets, a training set (typically 2/3rds to ¾s of the data), with the rest of the data forming a testing, or validation set. The training set is then sent through the neural network which iteratively improves the model using hidden layers. Briefly, these hidden layers assign weights to each explanatory parameter then compare the weighted model result to the testing set (the forward propagation process), the weights are then adjusted depending on the accuracy of the model (the backpropagation), and this process continues until the network is able to assign weights such that the network can predict the values correctly in most cases (Krenker *et al*. 2011). These hidden layers are termed thus because they estimate distributions that are not known explicitly using a combination of processes that are difficult, if not impossible, to visualise explicitly and thus they remain hidden – a “black box” process (Gevrey *et al*. 2006, Tan *et al*. 2015). ANNs have been utilised in ecology for a range of systems from mapping and identifying habitats using remote sensing (Brodick *et al*. 2019) to microhabitat temperature (Bryant & Shreeve 2002) but, as yet, there has been little advance in the field of population dynamics (Joseph 2020). In agriculture, using ANNs to forecast pest and disease outbreaks is a rapidly developing field (Junjing *et al*. 2019) and some (predominantly weather driven) forecast models have been developed for specific crops/locations such as coffee yield in Brazil (de Oliveira Aparecido *et al*. 2020), and for diseases such as powdery mildew in wheat (Agrawal & Mehti 2007), anthracnose crown rot in strawberries (Lu *et al*. 2017) and verticillium wilt in mint (Wheeler *et al*. 2019).

3.4.6.2 ANN Methods

To investigate the facility of ANNs for daily *B. aeneus* data, analysis was carried out using Keras (Chollet 2015), the high-level neural networks Application Programming Interface (API) for Tensorflow (Abadi *et al*. 2015) in R (R Core Team 2018). The model was trained on the meteorological, geographical and count parameters outlined in Table 3.1 with 60% of observations used as a training dataset, the remaining 40% as the testing dataset. The final network selection involved an Ordinal Regression Network with one hidden layer of 128 nodes. Greater numbers of layers and nodes were investigated but gave no better results and were increasingly heavy on computation time (measured in days). Ordinal Regression was selected as, due to the sparse nature of the counts (Fig. 3.6), the count data had to be apportioned into categories (0, 1-10, 11-30 and >30, selected to best represent the range of the data without having categories with too few data within them). Training sessions were run with a maximum of 100 epochs (iterations of the forward propagation process) with a patience level set at 10, to reduce overfitting in the modelling process (Zhou *et al*. 2020) i.e., the patience level controls when the model should report when it is complete; if the model runs 10 consecutive epochs with no significant improvement the training is deemed complete (Fig. 3.14).

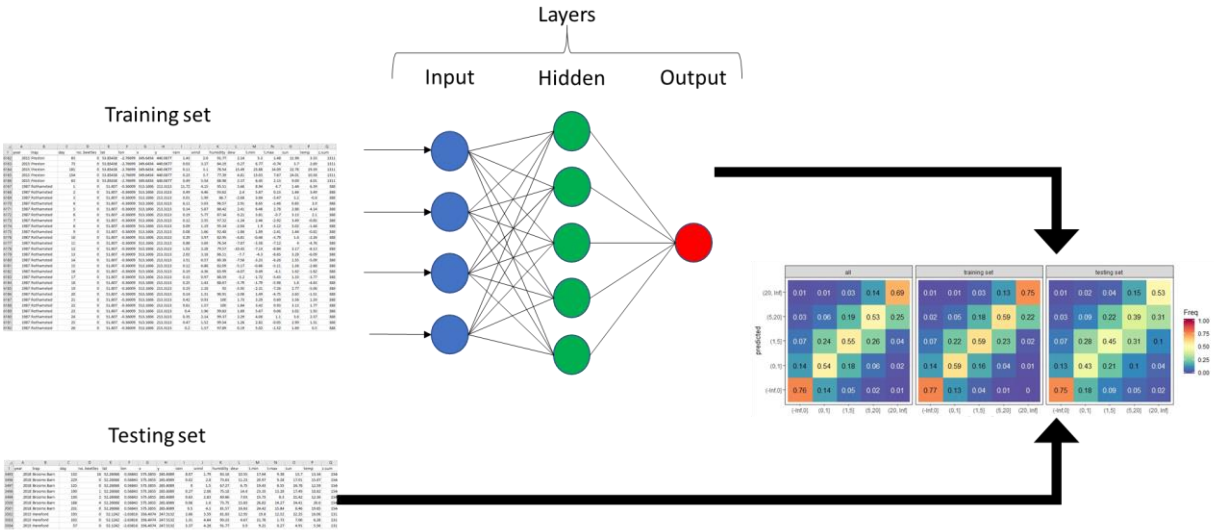


Fig. 3.13: Schematic of a neural network. The observational data are randomly assigned into two sets, a larger training set and a smaller testing set. The training set is then used as input into the ANN, within which are several interlinked nodes in one or more layers. Within these layers explanatory parameters are assigned weights which are fed forward through the ANN to a classification output, known as a confusion matrix. The matrix, and the assigned weights are then compared against the testing set. The comparison informs changes in weights through the hidden layers as alterations are back-propagated through them before being fed forward again, compared again and weights altered again. This iterative process continues until there is no improvement to the model weights or a pre-designated cut off point is reached.

The batch size (the number of samples from the training dataset used to estimate the error gradient – which is the statistical estimate of the difference between predicted and observed values after a model epoch) was set to 10, this is lower than has been recommended (e.g. Bengio 2012), but appears to provide a more stable and reliable training for the model, as noted by Masters & Luschi (2018) and also in the tuning of this model.

3.4.6.3 ANN results

Around 67% of zeroes are predicted correctly (69% in the testing set), however around 30% (29%-31%) observed zeroes are predicted in the 1-10 bin, with 2% forecast to be 10-30 beetles (Fig. 3.15). Conversely, as with the other models above, high counts prove difficult to predict, with only 0%-6% of high (>30) counts forecast accurately, and the majority (61%-69%) forecast to be 1-10 range, 21%-27% in the 11-30 range and as many as 6%-10% high counts predicted as zeroes (Fig. 3.15). The parameter selection behind these results are unknown due to the "black-box" nature of the hidden layer within the ANN (Gevrey *et al*. 2006, Tan *et al*. 2015). Hindcasts produced by the ANN to simulate suction-trap data from 2016 to 2018 (Fig. 3.16) appear to show that, in many cases the ANN is forecasting an extended season compared to what was observed, this is particularly true at sites such as Broom's Barn, Writtle and Rothamsted (Fig. 3.16). The ANN is also unable to forecast high counts well, especially at sites where high daily counts are rare, for example at Hereford (Fig. 3.16).

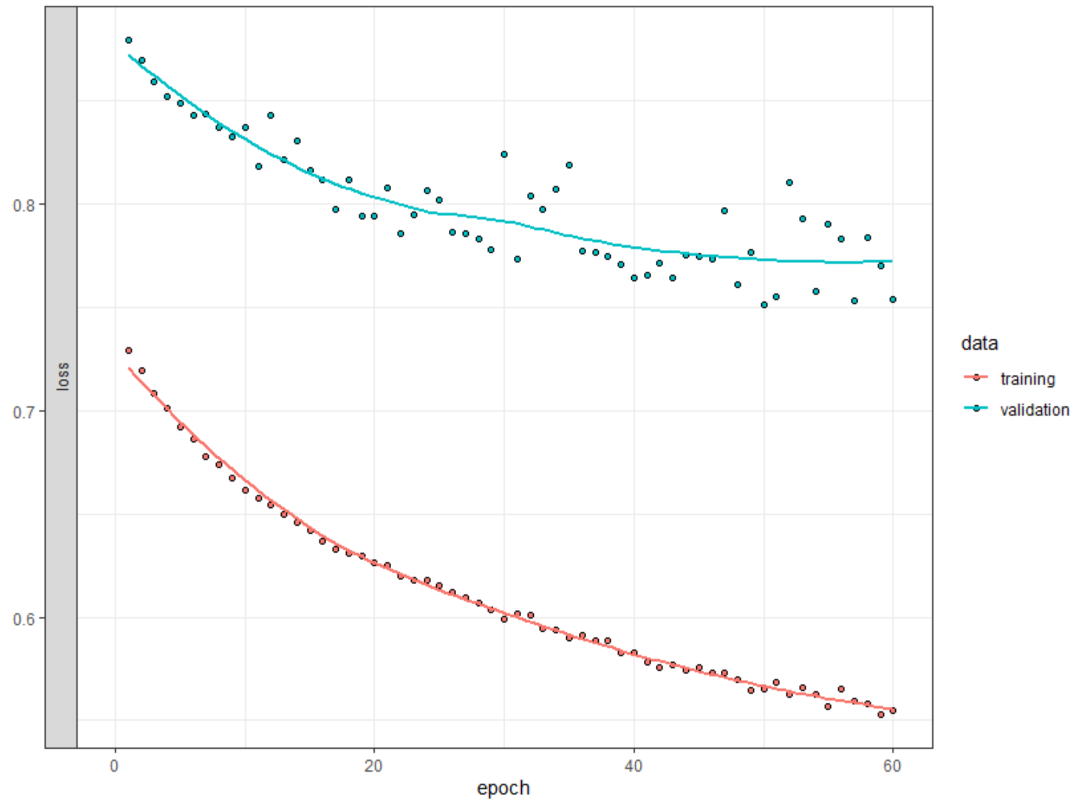


Fig. 3.14: The 'learning' performance of the Ordinal Regression Network. The validation curve indicates that, after around 20 epochs (iterations of the model) the model does not improve as the curvature becomes less steep and there are signs that over-fitting begins around 30 epochs, as the variance increases around the validation curve.

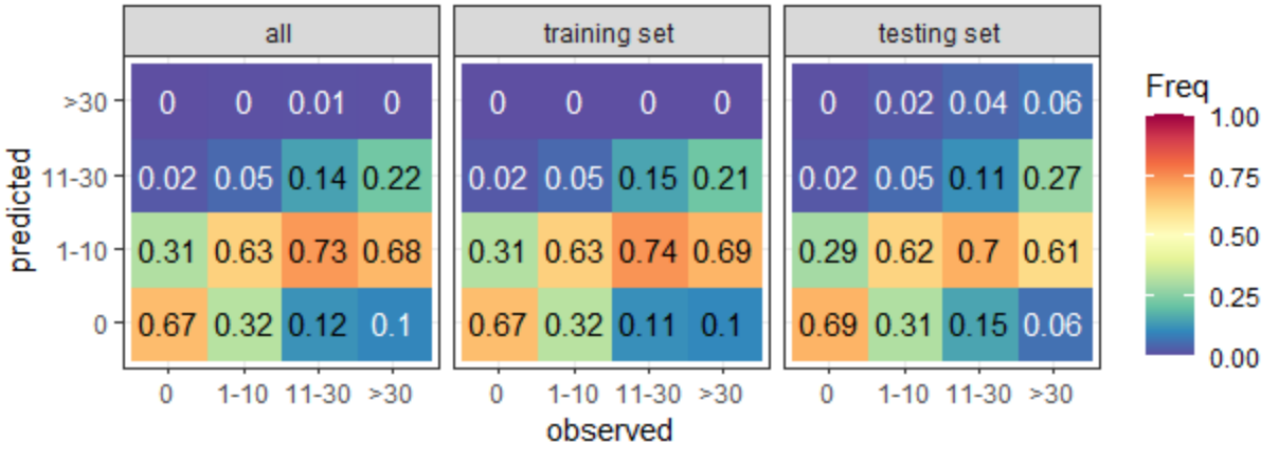


Fig. 3.15: Confusion matrix for the training dataset, the test dataset and both combined. For each of the 16 'bins' in each of the three matrices, the frequency value is scaled between 0-1 and coloured according to the level of agreement, with orange representing greater than 65% predictions assigned to that bin, conversely dark blue represents <25% of predictions in the bin (see legend). A perfect model would return 1 for each bin where observed & predicted values are the same (i.e. on a diagonal from bottom left to top right), with zeroes in other boxes.

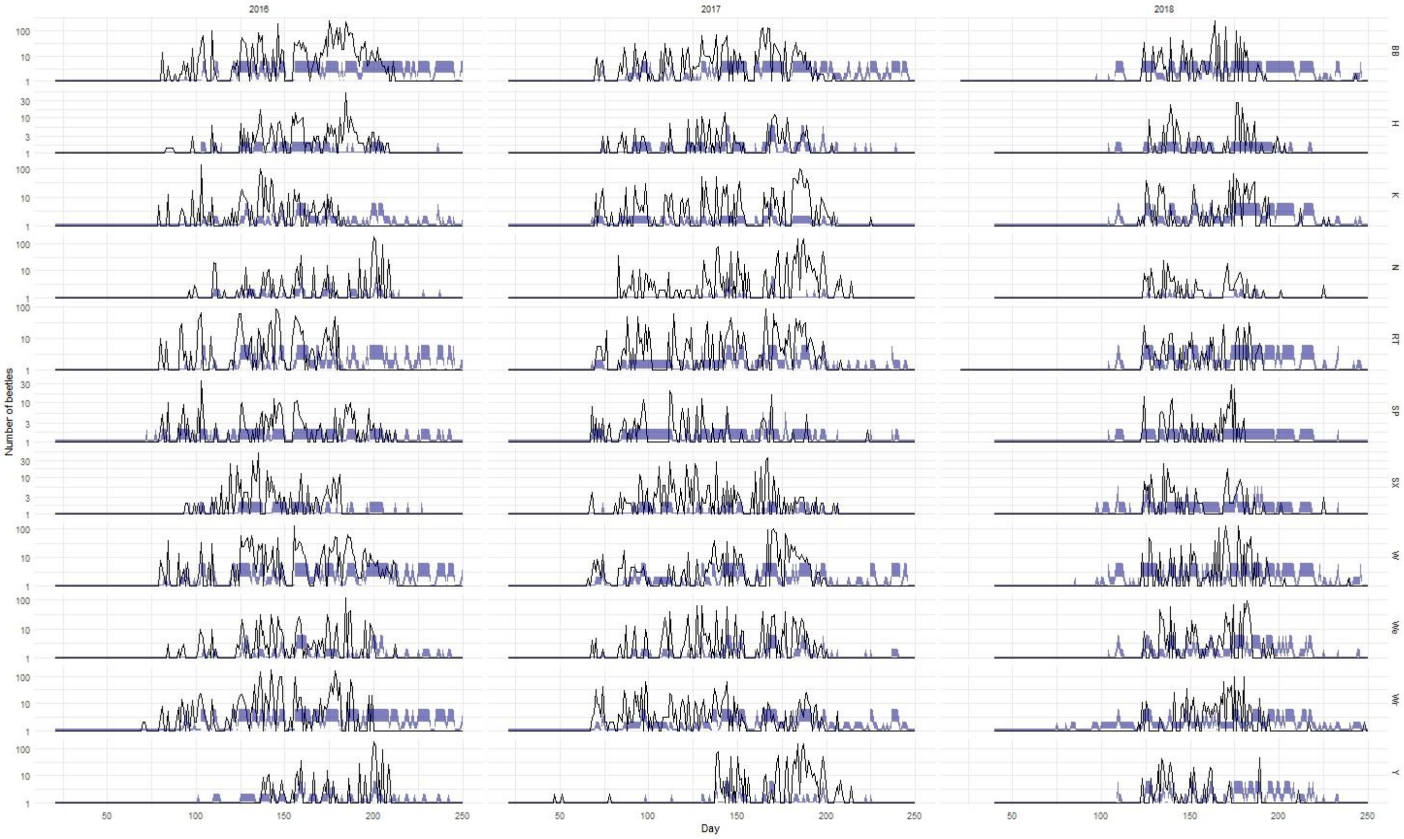


Fig. 3.16: A comparison between the predicted *Brassicogethes aeneus* count for each day (blue lines) and the observed values (black lines) for all sites 2016-2018 for every day between Julian Day 50 (19th February) and Julian Day 250 (31st July). Site abbreviations: BB: Broom's Barn; H: Hereford; K: Kirton; N: Newcastle; RT: Rothamsted; SP: Silwood Park; SX: Starcross; W: Wye; We: Wellesbourne; Wr: Writtle; Y: York.

3.5 Discussion

Currently, a forecast for *B. aeneus* phenology is available commercially (xarvio.com) though that model tool is now only available in a more comprehensive DSS package. Aside from the cost and accessibility implications of this change (Ryan 2019), it remains unable to forecast the scale of a *B. aeneus* migration event (Johnen *et al*. 2010, Ferguson *et al*. 2016). Knowledge of whether a migration event will be large or small could have a profound effect on the decision made at the farm level whether to take remedial control action or not. In addition, no models are available that forecasts migration into spring OSR crops at their susceptible stage (Johnen *et al*. 2010, Ferguson *et al*. 2016). The aim of this piece of research was to determine how best to model the daily data from the Rothamsted suction-trap network to inform a future DSS that extends the capabilities of current models. I present here four different methods for modelling this challenging dataset. The first, a GAM, failed to produce a satisfactory result, with the high number of zero counts proving too problematic to resolve. The other three methods produced varying results, none were able to predict high counts well, though lower counts and zero counts could be forecast.

The zero-inflated GLMM was the most parsimonious, with three parameters (sunshine, wind speed and day number) required for the binomial (presence-absence) step of the model and four parameters (number of beetles caught after June 1st the previous year, mean temperature, wind speed and longitude) required in the GLM step. These parameters were also utilised by the random forest model, with the exception of minimum temperature, which the random forest model selected instead of mean temperature. Minimum temperature was excluded in the model selection process for the zero-inflated GLM. The ANN did not outperform the zero-inflated GLM or the random forest model with the additional penalty of the “black box” nature of the model in that it is not known exactly how the hidden layer assigned weights to each parameter.

There is the possibility that beetle flight is driven by environmental cues not covered here, for example one possible trigger for beetle flight is the change in barometric pressure (Fournier *et al.* 2005, Pellegrino *et al.* 2013, Martini & Stelinski 2017, Martini *et al.* 2018). It would be a useful exercise to investigate how barometric pressure affects pollen beetle flight activity.

All models demonstrate that it is difficult to forecast rare extreme events – for example high catches – especially if those events are not previously recorded at a location. Given the rise in the *B. aeneus* population as measured by suction-traps (Chapter 2) this highlights a potential drawback in the utilisation of any models during a population expansion event. For *B. aeneus* it may be that population growth has levelled off in the south-east (Chapter 2) but if it continues to increase in areas where it is currently rarer then forecasts based on the model developed here would become rapidly more unreliable through underestimation. The difficulty of forecasting extreme events is accepted across disciplines (Stephenson *et al*. 2008, Goodwin & Wright 2010, Hitchens *et al*. 2013, Lerch *et al*. 2017, Farazmand & Sapsis 2019), including ecology (Denny *et al*. 2009, Van Doren & Horton 2018) and, whilst progress has been made there is not yet a satisfactory method to forecast outbreaks as extreme high counts are rare (Hitchens *et al*. 2013, Farazmand & Sapsis 2019).

It may be possible to overcome these limitations by using new techniques in frequentist statistics such as GLMM with bagging (Osawa *et al*. 2011) or combining the deep learning aspect of an ANN with, for example, state space modelling (Hyndman *et al*. 2008, Rangapuram *et al*. 2018), hybrid data mining (Guikema & Quiring 2012) or universal differential equations (Rackauckas *et al*. 2020). However, that step requires biostatical input and is beyond the scope of this thesis.

Chapter 4.

The spatial ecology of pollen beetles in Great Britain

Chapter Summary

Data on Brassicogethes aeneus *abundance as recorded in Rothamsted Insect Survey suction-traps and at field-level on yellow sticky-traps across Great Britain from 2008-2012 were used to investigate the spatial synchrony of British pollen beetles.*

Field data were loosely correlated with suction-trap data, however the sampling regimes in the two datasets were not fully compatible, with the sticky-trap data using asynchronous multi-day sampling that was hard to reconcile with daily or standardised weekly data from the suction-trap network.

Field data indicated that there was no spatial synchrony in B. aeneus *at the field level, meaning that it was impossible to predict beetle abundance in one field using data from a neighbouring field, indicating widespread heterogeneity at small scales. Over much larger landscape scales, suction-trap data showed some spatial synchrony, declining to zero at ~150 km.*

This lack of relationship between suction-trap data and field data, and the lack of synchrony detected in the field data raises doubts over the suitability of using suction-trap data to inform forecasts of migration into brassica crops.

4.1 Introduction

4.1.1 The need to consider spatial metrics in a DSS

Chapters two and three of this thesis demonstrate that there is some potential to estimate pollen beetle abundance and phenology from high-level movement of individuals as recorded in suction-traps. Potentially these data can provide useful information for a decision support system (DSS), however it is not yet demonstrated that these counts – observed or estimated – have any relationship to what occurs at the field level. It has long been known that the number of insects caught reduces with increasing height of a trap above ground level (Johnson 1957a, Taylor 1960) although there are differences between species (Wolfenbarger 1946, Taylor 1960). The pattern of insect abundance produces a density profile, the gradient of which is subject to weather (Johnson 1957a, Taylor 1960), particularly decreasing temperature (Johnson 1957b) and increasing wind speed with increasing height (Taylor 1960). For suction-traps, most research to date has focussed on the aphid fraction and much is known about their density profile (Johnson 1957a, Taylor 1960, 1973) and spatial distribution (Taylor 1973, Bell *et al*. 2012, 2015, 2019, Sheppard *et al*. 2016, Borhuis *et al*. 2020). Even so, whilst forecasting of aphid migrations from suction-traps has been in place for several decades (Cocu *et al*. 2005a, b, c, Harrington & Woiwod 2007, Bell *et al*. 2015), it is only relatively recently that these data have been shown to reflect aphid numbers within crops across Europe (McVean *et al*. 1999, Kasprowicz *et al*. 2008, Fabre *et al*. 2010). This broader understanding builds on earlier work by Leather & Lehti (1982) on the distribution patterns of just one aphid species (*Rhopalosiphum padi* (L.)) in Finland. No such analyses have been conducted for pollen beetles, and the question remains as to whether the suction-trap network is of use to growers as sentinels of infestation and outbreak at the field level. Specifically, to what degree do pollen beetle numbers in the suction-traps reflect numbers found in the field and at what spatial scale?

More generally, an understanding of the spatial dynamics of pest insects and the mechanisms that drive them is important in developing control strategies (Taylor 1986, Nestel *et al*., 2004, Cocu *et al*. 2005b, c). Spatial scaling becomes even more important when developing Integrated Pest Management (IPM) strategies (Sciarretta & Trematerra, 2014) as spatial patterns and amount of heterogeneity in pest populations have implications for control at both the field- (Park *et al*. 2007) and landscape scales (Hendrichs *et al*. 2007). DSSs, usually based on software models (Rose *et al*. 2016), are widely used in IPM to assist with crop management, including the timing and locations of control measures (Johnen *et al*. 2010, Sciaretta & Trematerra, 2014, Lindblom *et al*., 2017). Dynamic DSSs are those whose recommendations vary according to user inputs such as weather, soil, crop variables etc. (Rossi *et al*., 2014, Rose *et al*., 2016). Dynamic DSSs need to interpolate predictions across landscapes, perhaps even at farm level, to make them worthwhile and useful in practice (Rose *et al*., 2016, Lindblom *et al*., 2017, Lundström & Lindblom, 2018). As such, geostatistics – statistical methods for analysing spatial dependence among samples (autocorrelation and synchrony) and obtaining estimates for unsampled locations (Sciaretta & Trematerra, 2014) – can be powerful tools in elucidating pest systems (Liebhold & Gurevitch, 2002) and can improve the quality of DSSs for IPM.

To understand the role of spatially explicit models in DSSs, several themes are worthy of mention. Spatial autocorrelation, spatial synchrony and the relationship of the drivers of those systems, such as temperature and rainfall, often referred to under the Moran’s Theorem, all determine the accuracy and precision of any spatial prediction.

4.1.2 Spatial Ecology

The study of how space directly or indirectly affects ecological systems – termed spatial ecology (Tilman & Karieva 2018) – is a broad discipline that covers a range of separate, but related areas of study (Koenig 1999, Fletcher & Fortin 2018), including movement ecology focussed on dispersal and migration (Nathan *et al*. 2008, Fletcher & Fortin 2018). Across disciplines spatial ecology research is aimed at understanding the processes that affect species distribution and dynamics across space (Fletcher & Fortin 2019). One of the key challenges in spatial ecology is that the spatial structure of ecological populations is non-random, and observations are not independent of each other across space, a phenomenon known as spatial autocorrelation (Legendre 1993, Legendre *et al*. 2002, Dormann *et al*. 2007).

4.1.2.1 Spatial autocorrelation

Tobler's First Law states that "*everything is related to everything else, but near things are more related than distant things*" (Tobler 1970, Miller 2004). This truism is central to the understanding of spatial autocorrelation where populations at nearby locations are more similar (positive autocorrelation) or less similar (negative autocorrelation) than expected for randomly associated pairs of observations and are thus are not independent from each other (Legendre 1993, Dormann *et al*. 2007). Most ecological systems show some degree of autocorrelation across all spatial scales (Legendre & Fortin 1989, Legendre 1993) such that the distribution of individuals is neither uniform nor random (Legendre & Fortin 1989) and as such violate the assumptions of classical statistics (Legendre 1993, Keitt *et al*. 2002, Dormann *et al*. 2007). This may cause spurious associations within data if not accounted for (Lennon 2000, Keitt *et al*. 2002, Liebhold & Gurevitch 2002). However, discrete-time autocorrelation models (corAR1) were investigated by Bell *et al*.

(2020) when examining long term change in aphid populations using suction-trap data. They found that whilst temporal autocorrelation within model residuals was present, the positive autocorrelation detected was very weak. This is typical for many animal systems (Knape 2016).

4.1.2.2 Spatial synchrony

Spatial synchrony, defined as correlated fluctuations among populations that are separated in space (Hanski & Woiwod 1993, Liebhold *et al*. 2004, 2006, Walter *et al*. 2017) is a form of spatial autocorrelation among separate populations, both within and between species. In effect, it is a formalisation of the idea that spatial autocorrelation between populations can be a predictable result of known processes (Bjørnstad *et al*. 1999, Liebhold *et al*. 2004, Cocu *et al*. 2005b). Spatial synchrony has been demonstrated in a range of insect taxa, including Lepidoptera (Miller & Epstein 1986, Pollard 1991, Thomas 1991, Hanski & Woiwod 1993, Sutcliffe *et al*. 1996, Williams & Liebhold 2000, Peltonen *et al*. 2002, Raimondo *et al*. 2004, Cooke & Lorenzetti 2006, Haynes *et al*. 2019), Diptera (Choi *et al*. 2011, Santoiemma *et al*. 2019), Hymenoptera (Lantschner *et al*. 2019), Hemiptera (Hanski & Woiwod 1993, Estay *et al*. 2009, Bell *et al*. 2020) and Coleoptera (Peltonen *et al*. 2002, Økland *et al*. 2005, Aukema *et al*. 2006). Three primary drivers of spatial synchrony have been identified (Bjørnstad *et al*. 1999, Ripa 2000, Liebhold *et al*. 2004), and the effect of these drivers may differ at different time or spatial scales (Bjørnstad *et al*. 1999, Haynes *et al*. 2019): 1) Dispersal among populations causing a levelling effect between large and small populations as individuals spill-over from the former to the latter (Hanski & Woiwod 1993, Ranta *et al*. 1995, Heino *et al*. 1997, Kaitala & Ranta 1998, Palmqvist & Lundberg 1998). 2) Trophic interactions with other, synchronous predator/prey/parasite populations (Ims & Steen 1990, de Roos *et al*. 1991, 1998, Gurney *et al*. 1998, Bell *et al*. 2012). 3) The effects of widespread, spatially autocorrelated but density independent factors such as environmental fluctuations (Moran 1953, Royama 2012), named the Moran effect, or Moran's theorem (Royama 2005, 2012, Hansen *et al*. 2020).

4.1.2.3 Moran's theorem

Moran's theorem is named after Patrick Moran, whose seminal paper (Moran 1953) found that the synchrony of population dynamics across space had a simple mathematical relationship with spatially autocorrelated environmental drivers, usually temperature (Hansen *et al*. 2020). Being simple, Moran's theorem makes several assumptions about populations that are known to be difficult, or impossible to meet in field conditions (Lande *et al*. 1999, Engen & Sæther 2005, Liebhold *et al*. 2006, Hansen *et al*. 2020). Namely, the theory assumes: 1) No dispersal between populations, although Engen & Sæther (2005) assert that migration can be ignored when synchrony is examined at large spatial scales, provided that density regulation occurs, whereas Ranta *et al*. (1995) note that the effects of dispersal on synchrony can be detected by a theoretical decay in synchrony with increasing distance between populations. 2) A log-linear density dependence in populations. 3) Identical structure between population processes (Hansen *et al*. 2020). In addition, Kendall *et al*. (2000) found that it is difficult to separate the effects of environmental synchrony from other drivers such as dispersal or species interaction. Also, synchrony can change with time, along with patterns of spatial environmental autocorrelation (Koenig 2002, Post & Forchhammer 2002). There have been efforts to formulate more "generalised" Moran effect models, including those that relax the assumptions on log-linear density (Engen & Sæther 2005) or population structure (Liebhold *et al*. 2006), incorporate dispersal (Ripa 2000) or time scale specificity (Sheppard *et al*. 2016). Those points aside, the significance of Moran's theorem is that in principle, the cause of synchrony can be separate and independent from the drivers of cyclic population fluctuations (Bjørnstad & Grenfell 2001, Royama 2012), allowing some flexibility in determining the causes of population cycles and, independently, their synchrony (Ripa 2000, Royama 2012) and understanding the effects of wide-scale environmental disturbance (e.g. climate change) on populations (Sheppard *et al*. 2016, Hansen *et al*. 2020).

There are numerous ways to conceive spatial synchrony and thus test the evidence for Moran's theorem. For example, a spline corellogram model has been developed using nonparametric covariance functions (Bjørnstad *et al*. 1999, Bjørnstad & Falck 2001). In a spline correlogram analysis synchrony is measured by means of using the difference between successive observations to create a first-differenced (the series of changes in a population from one time period to the next) time series of log-abundance using the formula $\log N_t - \log N_{t-1} = z_t$, where N_t is the abundance or density at time *t* (Bjørnstad *et al*. 1999, Bjørnstad & Falck 2001). Doing this puts emphasis on the synchrony in population growth rates rather than raw abundance as the latter can lead to spurious correlation through adding weight to long-term trends that may arise for different reasons in different populations (Bjørnstad *et al*. 1999).

4.2 Aims and objectives

The main aims of this chapter are to quantify the spatial component in pollen beetle populations by linking suction-trap data with field collected data.

The objectives were;

1) to investigate the correlation between suction-trap and field data to determine the geographical scale at which trap counts can be interpreted

2) to determine measures of temporal and spatial autocorrelation

3) to determine the viability of using suction-trap data in farmer Decision Support System models to forecast immigration into oilseed rape fields

4.3 Materials and Methods

4.3.1 Temporal and spatial patterns in field abundance

4.3.1.1 Sticky-trap data collection and standardisation

Pollen beetle immigration from late February – early May 2008 to 2012 for a total of 138 winter OSR field sites was monitored using yellow sticky-traps (Cook *et al*. 2014, Skellern *et al*. 2017). At each site, standard "wetstick" 10 x 20 cm sticky traps (Oecos, Kimpton, Hertfordshire) were set, clipped to extendable poles at a 45° angle to the vertical (Blight & Smart, 1999) and maintained at crop canopy height through the trapping period. Traps were placed 3 m into the crop from the field edge, oriented to face outwards from the crop. Where practicable, traps were changed every 3-4 days, but this was often not the case with shorter and longer time periods common (2-21 days, with ~80% of intervals <7 days) (Skellern *et al*. 2017). Traps were operated by volunteer farmers or crop consultants at fields across the main UK OSR cropping region (Fig. 4.1) except for sites at Rothamsted and Woburn which were operated by Rothamsted staff. Some sites, including Rothamsted and Woburn, ran traps in multiple fields in the same year, or in individual fields in multiple years. Collected traps were returned to Rothamsted for assessment of pollen beetle catch. Due to the irregular, and largely asynchronous trapping periods (Fig. 4.2), trap data were aggregated to weekly counts across standardised weeks (16th February to 25th May). The Central England Temperature index (Parker *et al*., 1992) for each trapping day was analysed and any days below a maximum temperature threshold of 11°C were excluded, 11°C was set as the threshold as this approximated to the temperature threshold for an increased chance of a non-zero count (see Chapter 3, Fig. 3.12) with a buffer of ~1°C. Trap samples were then averaged across the remaining days and summed into standardised weeks.

4.3.1.2 Suction-trap data collection and standardisation

The Rothamsted Insect Survey (RIS) network of suction-traps (Macaulay *et al*. 1988, Bell *et al*. 2015) began in 1964, primarily to monitor aphids (Harrington & Woiwod 2007). The traps sample aerial insect populations at 12.2 m above ground level at a rate of 50 m^3 min^{-1}(Macaulay *et al*. 1988). Daily samples are taken during the aphid season (roughly late March to November), with weekly samples over winter and early spring. The aphids are removed, identified and counted and the bycatch is stored in a mixture of ethanol & glycerol (Shortall 2013). The presence of glycerol means that, though samples do become dehydrated over time, they remain undamaged on rehydration. The bycatch was investigated for pollen beetles, with all *B. aeneus* present identified and counted. Data from all suction-trap sites from 2015-2018 were used, daily or weekly samples (depending on the time of year). See Chapters 2 and 3 for further details of the suction-trap sampling. The data from the traps were standardised to the same weeks as the field data with daily samples assigned to the relevant week and weekly samples divided between weeks in the same way as the sticky-trap data, i.e. days with a maximum temperature below 11°C were excluded and the count was then divided between the remaining days (Fig. 4.2).

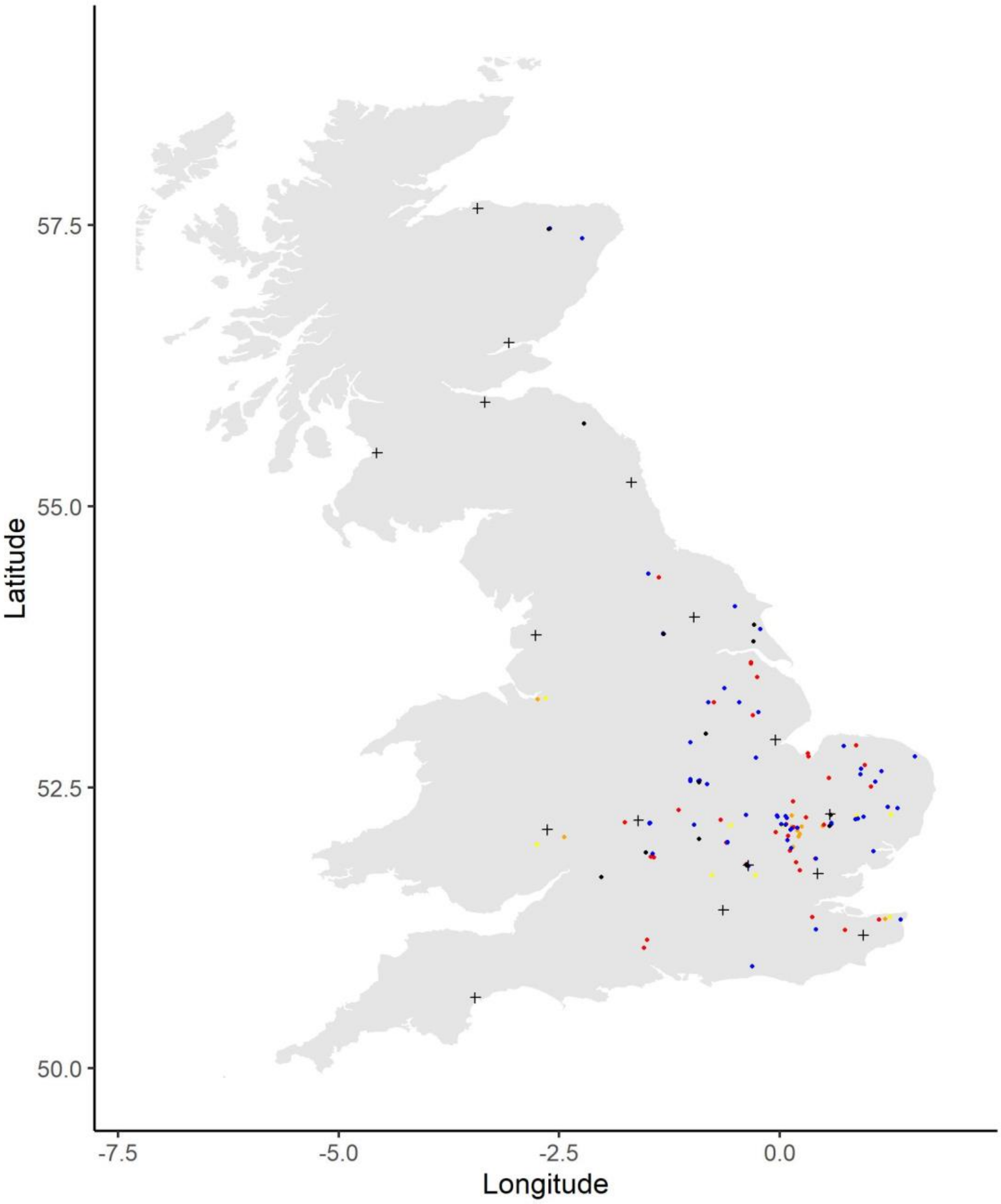
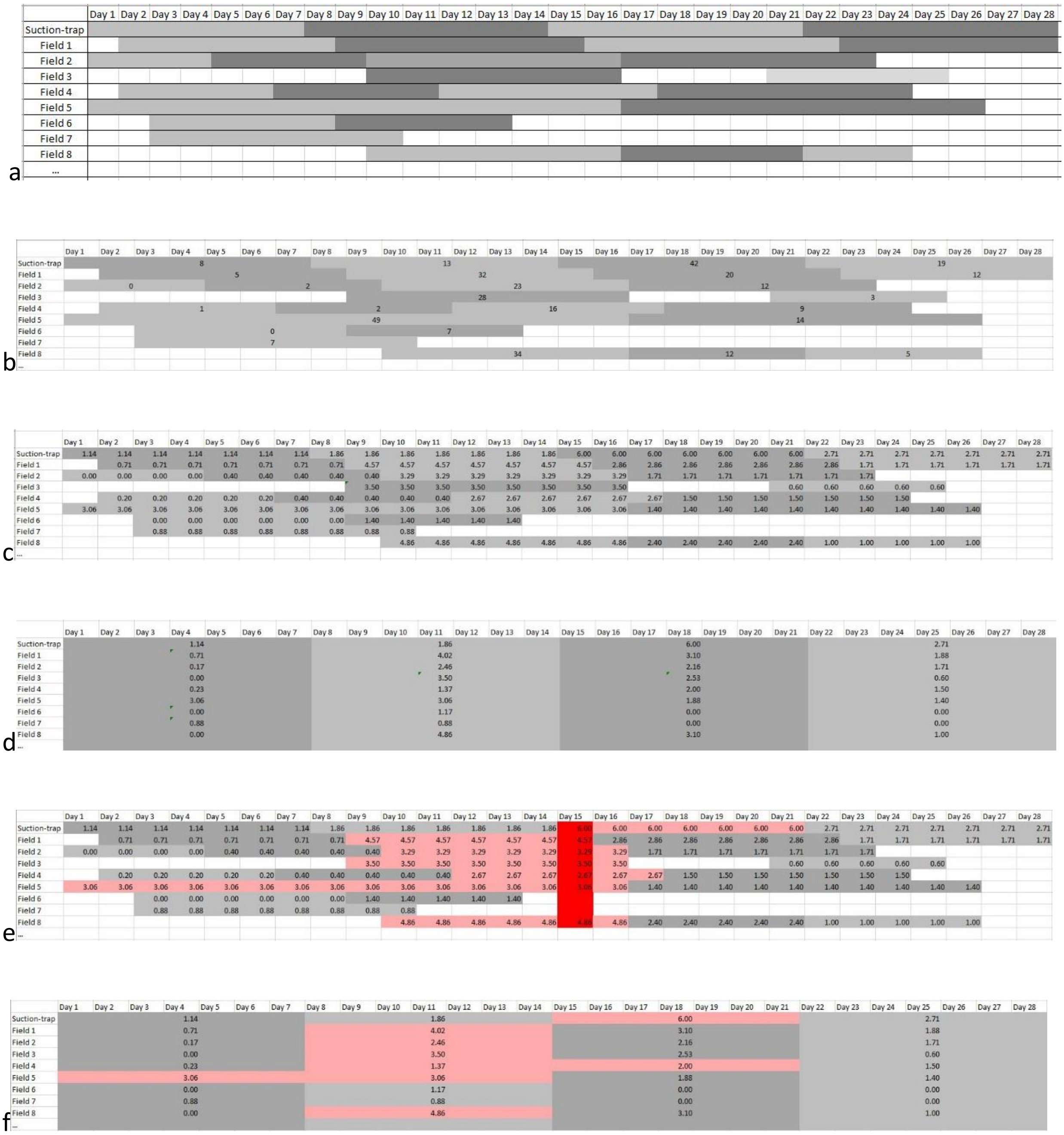


Fig. 4.1: Locations of Rothamsted Insect Survey suction-trap sites (crosses) and sticky trap field sites (points). Field sites are coloured by year: Yellow = 2008; Orange = 2009; Red = 2010; Blue = 2011; Black = 2012.



	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7	Day 8	Day 9	Day 10	Day 11	Day 12	Day 13	Day 14	Day 15	Day 16	Day 17	Day 18	Day 19	Day 20	Day 21	Day 22	Day 23	Day 24	Day 25	Day 26	Day 27	Day 28
Suction-trap																												
Field 1																												
Field 2																												
Field 3																												
Field 4																												
Field 5																												
Field 6																												
Field 7																												
Field 8																												
...																												

	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7	Day 8	Day 9	Day 10	Day 11	Day 12	Day 13	Day 14	Day 15	Day 16	Day 17	Day 18	Day 19	Day 20	Day 21	Day 22	Day 23	Day 24	Day 25	Day 26	Day 27	Day 28
Suction-trap				8							13							42							19			
Field 1					5							32							20							12		
Field 2		0					2						23							12								
Field 3												28											3					
Field 4				1					2					16							9							
Field 5									49												14							
Field 6						0					7																	
Field 7						7																						
Field 8													34						12					5				
...																												

	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7	Day 8	Day 9	Day 10	Day 11	Day 12	Day 13	Day 14	Day 15	Day 16	Day 17	Day 18	Day 19	Day 20	Day 21	Day 22	Day 23	Day 24	Day 25	Day 26	Day 27	Day 28
Suction-trap	1.14	1.14	1.14	1.14	1.14	1.14	1.14	1.86	1.86	1.86	1.86	1.86	1.86	1.86	6.00	6.00	6.00	6.00	6.00	6.00	6.00	2.71	2.71	2.71	2.71	2.71	2.71	2.71
Field 1		0.71	0.71	0.71	0.71	0.71	0.71	0.71	4.57	4.57	4.57	4.57	4.57	4.57	4.57	2.86	2.86	2.86	2.86	2.86	2.86	2.86	1.71	1.71	1.71	1.71	1.71	1.71
Field 2	0.00	0.00	0.00	0.00	0.40	0.40	0.40	0.40	0.40	3.29	3.29	3.29	3.29	3.29	3.29	3.29	1.71	1.71	1.71	1.71	1.71	1.71	1.71					
Field 3									3.50	3.50	3.50	3.50	3.50	3.50	3.50	3.50					0.60	0.60	0.60	0.60	0.60			
Field 4		0.20	0.20	0.20	0.20	0.20	0.40	0.40	0.40	0.40	0.40	2.67	2.67	2.67	2.67	2.67	2.67	1.50	1.50	1.50	1.50	1.50	1.50	1.50				
Field 5	3.06	3.06	3.06	3.06	3.06	3.06	3.06	3.06	3.06	3.06	3.06	3.06	3.06	3.06	3.06	3.06	1.40	1.40	1.40	1.40	1.40	1.40	1.40	1.40	1.40	1.40		
Field 6			0.00	0.00	0.00	0.00	0.00	0.00	1.40	1.40	1.40	1.40	1.40															
Field 7			0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88																		
Field 8										4.86	4.86	4.86	4.86	4.86	4.86	4.86	2.40	2.40	2.40	2.40	2.40	1.00	1.00	1.00	1.00	1.00		
...																												

	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7	Day 8	Day 9	Day 10	Day 11	Day 12	Day 13	Day 14	Day 15	Day 16	Day 17	Day 18	Day 19	Day 20	Day 21	Day 22	Day 23	Day 24	Day 25	Day 26	Day 27	Day 28
Suction-trap				1.14							1.86							6.00							2.71			
Field 1				0.71							4.02							3.10							1.88			
Field 2				0.17							2.46							2.16							1.71			
Field 3				0.00							3.50							2.53							0.60			
Field 4				0.23							1.37							2.00							1.50			
Field 5				3.06							3.06							1.88							1.40			
Field 6				0.00							1.17							0.00							0.00			
Field 7				0.88							0.88							0.00							0.00			
Field 8				0.00							4.86							3.10							1.00			
...																												

	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7	Day 8	Day 9	Day 10	Day 11	Day 12	Day 13	Day 14	Day 15	Day 16	Day 17	Day 18	Day 19	Day 20	Day 21	Day 22	Day 23	Day 24	Day 25	Day 26	Day 27	Day 28
Suction-trap	1.14	1.14	1.14	1.14	1.14	1.14	1.14	1.86	1.86	1.86	1.86	1.86	1.86	1.86	6.00	6.00	6.00	6.00	6.00	6.00	6.00	2.71	2.71	2.71	2.71	2.71	2.71	2.71
Field 1		0.71	0.71	0.71	0.71	0.71	0.71	0.71	4.57	4.57	4.57	4.57	4.57	4.57	4.57	2.86	2.86	2.86	2.86	2.86	2.86	2.86	1.71	1.71	1.71	1.71	1.71	1.71
Field 2	0.00	0.00	0.00	0.00	0.40	0.40	0.40	0.40	0.40	3.29	3.29	3.29	3.29	3.29	3.29	3.29	1.71	1.71	1.71	1.71	1.71	1.71	1.71					
Field 3									3.50	3.50	3.50	3.50	3.50	3.50	3.50	3.50					0.60	0.60	0.60	0.60	0.60			
Field 4		0.20	0.20	0.20	0.20	0.20	0.40	0.40	0.40	0.40	0.40	2.67	2.67	2.67	2.67	2.67	2.67	1.50	1.50	1.50	1.50	1.50	1.50	1.50				
Field 5	3.06	3.06	3.06	3.06	3.06	3.06	3.06	3.06	3.06	3.06	3.06	3.06	3.06	3.06	3.06	3.06	1.40	1.40	1.40	1.40	1.40	1.40	1.40	1.40	1.40	1.40		
Field 6			0.00	0.00	0.00	0.00	0.00	0.00	1.40	1.40	1.40	1.40	1.40															
Field 7			0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88																		
Field 8										4.86	4.86	4.86	4.86	4.86	4.86	4.86	2.40	2.40	2.40	2.40	2.40	1.00	1.00	1.00	1.00	1.00		
...																												

	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7	Day 8	Day 9	Day 10	Day 11	Day 12	Day 13	Day 14	Day 15	Day 16	Day 17	Day 18	Day 19	Day 20	Day 21	Day 22	Day 23	Day 24	Day 25	Day 26	Day 27	Day 28
Suction-trap				1.14							1.86							6.00							2.71			
Field 1				0.71							4.02							3.10							1.88			
Field 2				0.17							2.46							2.16							1.71			
Field 3				0.00							3.50							2.53							0.60			
Field 4				0.23							1.37							2.00							1.50			
Field 5				3.06							3.06							1.88							1.40			
Field 6				0.00							1.17							0.00							0.00			
Field 7				0.88							0.88							0.00							0.00			
Field 8				0.00							4.86							3.10							1.00			
...																												

Fig. 4.2: Schematic demonstrating the challenge of asynchronous weekly sampling. Consider a hypothetical situation with a Suction-trap (ST) & eight field sites, the ST runs weekly and over the 28 day period is changed four times whilst the field sites are monitored for varying time periods - shown in (a) The number of beetles caught in those sampling periods is shown in (b). We could average these catches across the days to give a standardised weekly count (c, d). However if we have a single warm day that potentially encourages beetle flight (Day 15 in this example, coloured red with affected sampling periods highlighted in pink) (e) and we average across the sampling periods (f) then the beetles flying on that day would contribute more to the average sample of the previous week (or weeks) than the current week at all bar one of the field sites.

4.3.1.3 Analysis

All data were analysed using the R statistical package (R Core Team 2018). Field data were compared to suction-trap data from Rothamsted by Pearson correlation. The resulting annual correlation estimates were then plotted against distance from Rothamsted.

4.3.2 Temporal and spatial patterns of field and suction-trap data

4.3.2.1 Visualising and modelling spatial autocorrelation

Of the 15 standard weeks assessed the first week and the last three weeks were zero-inflated and were not analysed. The remaining weeks 2-12 (23 Feb – 03 May) were used for analysis. After checking the statistical distribution of the standardised weekly data from both datasets using the R library fitdistrplus (Delignette-Muller & Dutang, 2015), a generalized additive model (GAM) from the R package mgcv (Wood 2017) with location as a smooth term was used to produce a general trend that was interpolated across the spatial extent of the suction-trap and sticky-trap networks (Fig. 4.1). Due to the sparse nature of the data each standard week across all years was analysed separately alongside a combined total for all the data.

4.3.2.2 Quantifying spatial synchrony: spline correlogram analysis

To assess spatial synchrony in the field data 2008-12 and the suction-trap data 2014-18 for all sites (this thesis, Chapter 3) multivariate spline correlograms were used, where the y intercept indicated the local covariance function (Moran's I), the x estimated the spatial extent (km), and, the intercept of the spline at x gave an estimate of the correlation length. A bootstrap confidence interval was calculated by sampling with replacement among sites (deleting cross-correlations arising from comparing a site with itself). One thousand bootstrap iterations were done, and the confidence intervals calculated using the percentile method (Efron & Tibshirani, 1993). Spline correlograms were calculated using the ncf package in R (Bjørnstad & Bjørnstad, 2016).

4.4 Results

4.4.1 Comparing field and suction-trap data

The relationship between field and suction-trap data at a weekly level generally shows a moderate to strong relationship, although there is a lot of scatter (Fig. 4.3). It should be noted that for some locations correlation remains quite strong even at a distance, and the effect of year appears stronger than that of distance. Many sites showed no, or negative correlation with the suction-trap data in some years, which may be due to the effect of standardising asynchronous trapping periods (Fig. 4.2). As noted in Chapter 3, the daily counts of beetles from the suction-traps can vary by orders of magnitude from day to day, which may cause spurious results when comparing two different traps (as demonstrated in Fig 4.2).

4.4.2 Generalized Additive Model of field data

The GAM maps (Fig. 4.4) demonstrate that there is no clear national trend in the week to week abundance of pollen beetles as measured in the field. In all, only three weeks (7, 8 and 10) and the combined dataset showed any significant spatial correlation. It is also notable that the deviance explained was extremely low (<10%) in all cases apart from week 8 (Table 4.1). Generally, the counts were highest in south-east however the changes in site distribution and the variation in counts make it difficult to determine whether the direction of the trend (Fig. 4.4) is a reduction in abundance towards the north (weeks 2, 11, 12), the north-east (weeks 6, 7, 10), the east (weeks 3, 5, 9) or the south-east (weeks 4, 8). The rate of change across the interpolated values is also variable (Fig. 4.4), ranging from steep changes in abundance reflected by tight isoclines in some weeks (e.g. weeks 2 & 8) to less steep in others (e.g. weeks 4 & 6).

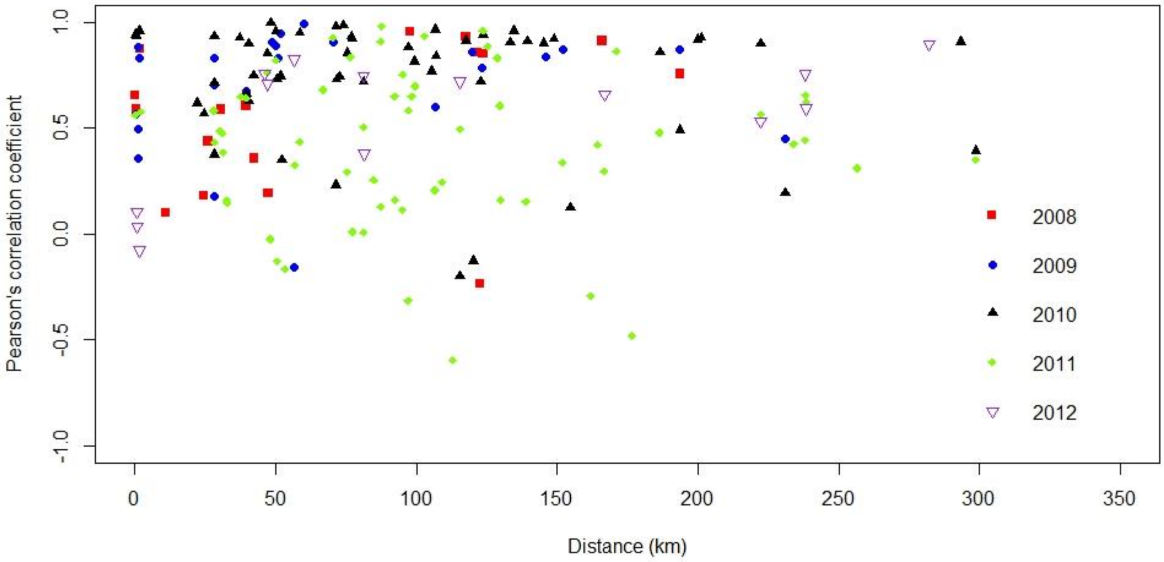


Fig. 4.3: Pearson's correlation coefficient for annual comparisons of standardised weekly *Brassicogethes aeneus* populations in field sites (sticky-traps) to the Rothamsted suction-trap plotted against the distance of each site from the suction-trap in kilometres. Different years are plotted in different colour/shape combinations (see legend). Two field sites: The Beeches (2011) and Haddo (2008 & 2011) both more than 600 kms away from the suction-trap have been excluded from this plot for clarity.

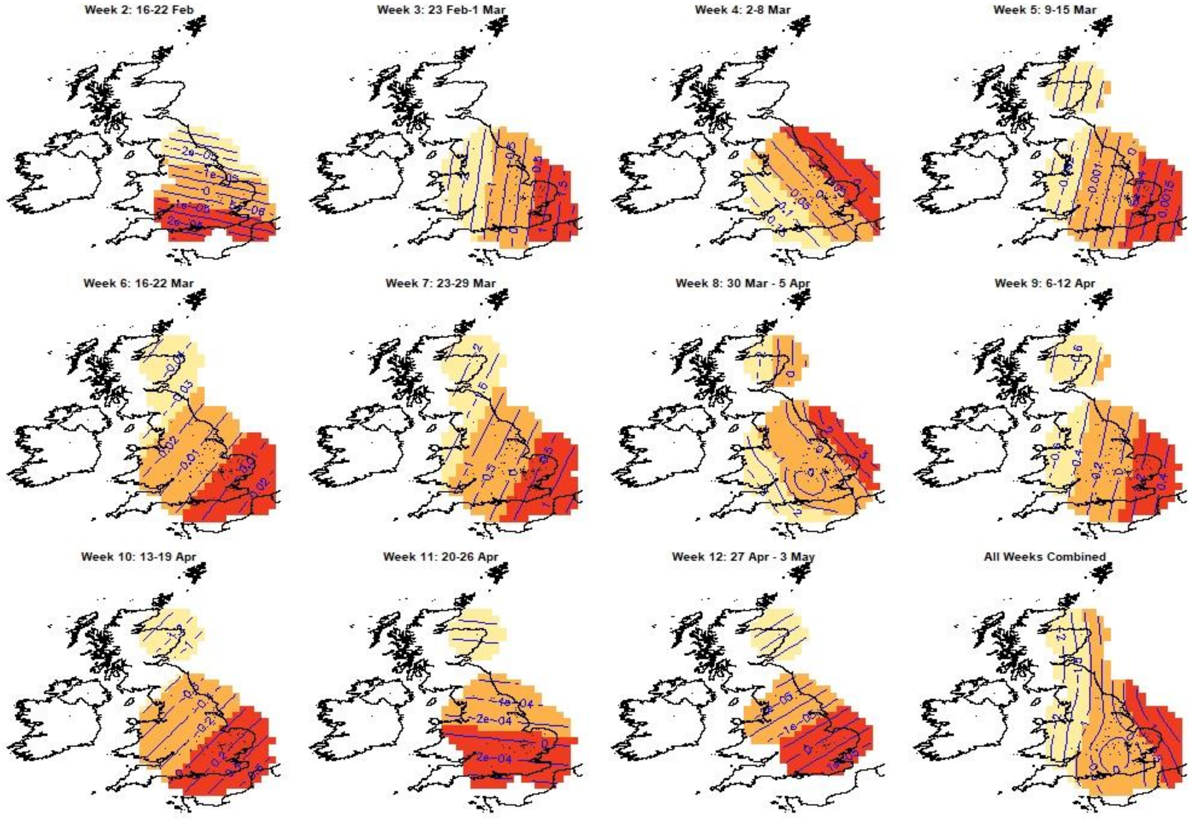


Fig. 4.4: Surface plots for GAM of sticky-trap and suction-trap data for individual weeks and the full dataset combined (bottom right). Red indicates higher counts, yellow lower counts. The isoclines on the maps are deviations from the GAM intercept in abundance. The unit value of each set of isoclines vary for each model from very small changes over space (e.g. week 12) to large differences (e.g. all weeks combined). The general pattern is for higher counts in the east and/or south-east of England, although the direction and amplitude of the drop from higher counts to lower changes from week to week.

4.4.3 Multivariate spline analysis of field data

The lack of pattern demonstrated above by the generalized additive models is reinforced by the multivariate spline correlogram of the field sticky-trap sites (Fig. 4.5). Whilst the spline intercepts the zero line at just above 122 km, from zero distance onwards the lower confidence interval crosses the zero-line indicating that there is no correlation even at a distance of zero. There is some evidence of correlation between sites becoming negative at greater distance, however the confidence intervals illustrate that no substantial inference should be drawn from this.

Week no.	No. of sites	edf	Deviance explained (%)	P
2	106	0.0000226	0.000243	0.502
3	44	1.114	8.14	0.16
4	136	0.2729	0.415	0.334
5	168	0.007042	0.00405	0.545
6	175	0.8611	0.0858	0.33
7	181	1.735	7.48	<0.001***
8	168	12.74	23.6	<0.001***
9	149	1.203	2.41	0.0506
10	107	1.486	4.92	0.00601**
11	76	0.0022	0.00353	0.463
12	51	0.0002384	0.000054	1
Combined	189	8.965	17.1	<0.001***

Table 4.1: Summary statistics for Generalized Additive Model shown in Fig. 4.4, edf is Estimated Degrees of Freedom which is estimated within the model.

4.4.4 Multivariate spline analysis of suction-trap data

Conversely, the spline correlogram for the suction-traps (Fig. 4.6) shows positive correlation between sites to 204.7 km (though the lower bound of the confidence interval suggests that ~150 km is more prudent). There is good evidence that synchrony declines with distance, from a Moran's index value of 0.23.

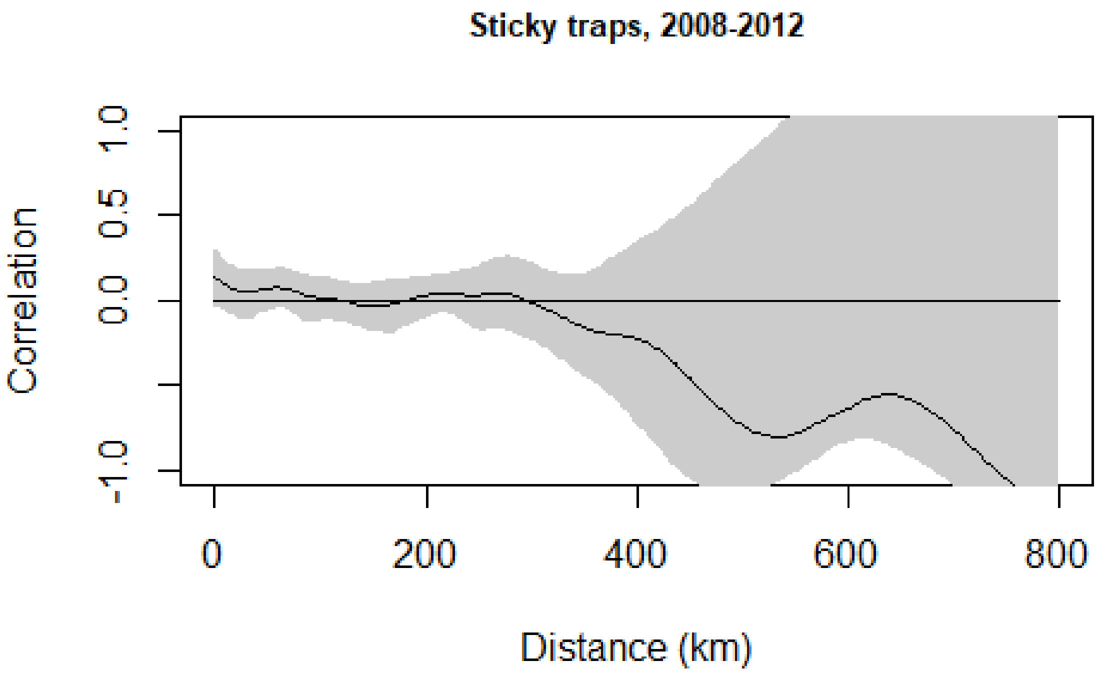


Fig. 4.5: Spline correlogram of standardised weekly data for all field sites (2008-2012).

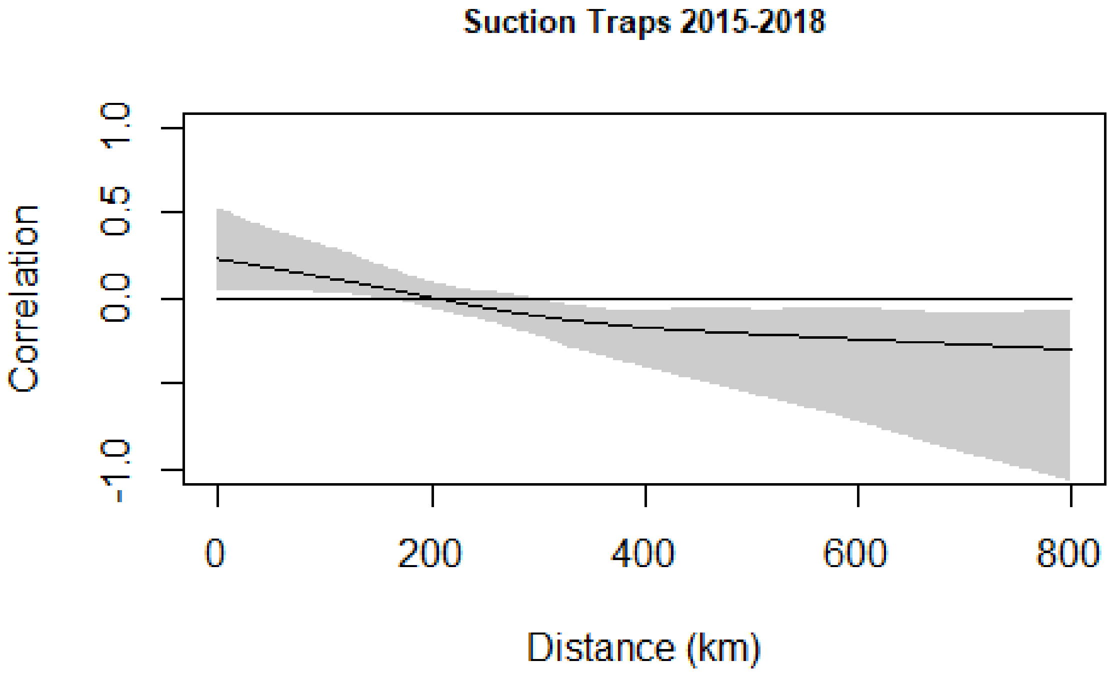


Fig. 4.6: Spline correlogram of daily data for all suction-traps (2015-2018).

4.5 Discussion

4.5.1 Discussion of the findings

There is very little spatial relationship between counts of *Brassicogethes aeneus* populations at different sites over time, instead they appear locally eruptive and independent. There is at best a positive spatial synchrony across the number of beetles caught in flight by the suction-trap network at a distance of up to around 150 km, beyond which synchrony breaks down. For field sites, even a weak relationship in spatial synchrony is absent suggesting that *B. aeneus* population dynamics may amount to a chaotic system – or at least display complex nonlinear dynamics that approximate to chaos (Turchin & Taylor 1992, Solé & Bascompte 1995, Zhou *et al*. 1997, Dennis *et al*. 2001). However, the data are not of sufficient quality to investigate this further (Schaffer 1985). Chaotic or not, it appears from the results here that it is impossible to forecast the population numbers at the weekly level from one field to the next using sticky-trap data, which is expected of chaotic systems (Solé & Bascompte 1995). This finding reinforces the results of Skellern *et al*. (2017) who found that, whilst meteorological parameters and the area of OSR grown in the previous season were predictors of pollen beetle numbers at the field scale, other landscape factors such as area of woodland and length of hedgerow local to the site showed no relationship to counts. However, it is unlikely that the system is truly chaotic as the suction-trap data do demonstrate that there is some, albeit very weak, spatial correlation among suction-traps.

The lack of a spatial signal or pattern contrasts with Chapters 2 and 3, that demonstrated that there is some potential to forecast temporal abundance – either from populations in the previous year/generation (Chapters 2, 3) or from meteorological and location data (Chapter 3). This could be due, in part, to the nature of the analyses in this chapter in that the comparison of data between traps on a standardised week-by-week basis across a latitudinal gradient may fail to account for small-scale, phenologically-driven outbreaks, whereas the analyses in the previous chapters are annual or daily assessments at a site-specific level, albeit across several sites. It is also probably the case that the datasets presented here are not long enough at 5 years for the sticky-trap data and 4 years for the suction-traps for comprehensive and successful analysis. Schaffer (1985) suggests that to detect low dimensional chaos in an ecological system requires i) a continuous sampling period that has ii) a frequency not less than ~10 data points per orbit (a complete rotation through the phase space of the system (Shaw 1981) – analogous to a population cycle) and iii) a sampling period long enough to encompass 30-50 orbits/cycles. Nearly all ecological research will fail to provide enough data under these conditions, but it was possible, even with this limited data, to detect the presence of spatial synchrony in the suction-trap data. Further investigation into long-term samples at other suction-traps may provide enough spatial coverage to better understand the synchrony (or lack of) of *B. aeneus* population dynamics.

4.5.2 Other studies that show a lack of spatial synchrony

Studies reporting a lack of spatial synchrony in populations are rare and often caveated due to the expectation that the Moran effect or other drivers will push systems towards synchrony more often than not in support of Tobler's First Law. Wool (2002) found a lack of synchrony in pistachio aphid *Baizongia pistaciae* (L.) gall abundance over 20 years in Israel. This was attributed to synchronising drivers working on other, unstudied, aspects of the aphid's complex life cycle. Processes affecting the gall-forming stage may be working in different directions to those affecting the root-feeding and migratory stages of the cycle (Yukawa 2000, Wool 2002) and thus synchrony may be present, undetected, in the pistachio aphid (Wool 2002). In the North Sea copepod *Calanus finmarchicus* (Gunnurus) a loss in spatial synchrony was attributed to the effect of climate change on populations, with other plankton species retaining, or even showing an increase in synchrony (Defriez *et al*. 2016). Bell *et al*. 2021 (in prep.) have found similar evidence for a lack of spatial synchrony in peach-potato aphids *Myzus persicae* (Sulzer) recorded across a network of water traps in East Anglia. These data agree well with this pollen beetle study in that both support the theory that these field-level traps detect locally eruptive dynamics that are scaled at no more than a few hundred metres. Indeed, the lack of correlation with other locally sited traps, even at farm level, suggests that an effective DSS would require such a large field-level trap campaign aiming to include many traps per farm and many farms per landscape, that it may prove unsustainable.

4.5.3 Spatial relationships in suction-trapped aphids

This research finds that, in the suction-traps there is spatial synchrony in *B. aeneus* with a maximum covariance as measured by Moran's I of 2.3, declining to zero at around 150 m, however field level data showed no synchrony, indicating that it may not be possible to show a link between field counts and data from suction-traps. In aphids, the relationship between suction-trap data and field counts has been established, if only in a limited fashion. Leather & Lehti (1982) first reported a relationship between suction-trap catches and field populations of *Rhopalosiphum padi* in Finland. There, significant correlations exist between alatae (winged, migrating aphids) caught in the trap and the logarithm of the number of aphids per 100 tillers in several cereal crops and also the number of gynoparae (parthenogenetic female aphids) found on the primary host *Prunus padus* L., albeit at a trap height of 1.2 m (Leather & Lehti 1982). McVean *et al*. (1999) found a regional relationship between counts of the pea aphid *Acyrthosiphon pisum* Mordvilko in catches from 12.2 m suction-traps and both crop area and aphid density within pea *Pisum sativum* L. fields at the time when the crops would be most susceptible to aphid damage. They suggested that variability in this relationship may be due to the lag between records of aphid migration in suction-traps and the establishment of populations in the field (McVean *et al*. 1999). These studies used relatively simple comparative statistics and it was not until Cocu *et al*. (2005a, b, c) that geostatistical methods were brought to bear on suction-trapped aphid data. This series of papers established the geographical pattern of *M. persicae* in suction-traps across Europe using the SADIE (spatial analysis by distance) technique (Cocu *et al*. 2005a), and used spatial analysis (Cocu 2005b) and multiple linear regression (Cocu 2005c) to show a significant association between landscape variables, particularly area of oilseed rape sown, and climate variables and the pattern of distribution in suction-traps, but did not attempt to compare suction-trap and field data. An indication that suction-trap catches reflect field data at the clonal level was noted by Kasprowicz *et al*. (2008) who observed that the more abundant clones of *M. persicae* in Scotland were found in similar proportions in field-caught and suction-trapped aphids, but did not produce a formal test of these relationships. The first, and thus far only, demonstration that suction-trap catches can predict numbers of aphids in the field was provided by Fabre *et al*. (2010) who used a hierarchical Bayesian modelling framework to show that suction-trap catches could predict to a satisfactory level the number of alate *R. padi* found in fields in northern France. The sparseness of the research in this field, even in such an intensively studied system as aphids in the suction-trap network (Bell *et al*. 2015) demonstrates the difficulty in determining any relationship between suction-trap sampling and field sampling.

4.5.4 Spline correlogram studies that show spatial synchrony

Although no spatial synchrony was detected from the sticky-traps at crop height in this study, the suction-trap data did show some synchrony which decayed with distance. This finding is reflected by Bell *et al*. (2020) who found evidence of spatial synchrony in suction-trapped aphids at

12.2 m, but not moths caught in light-traps at 1.2 m. The comparison with these data should be caveated by noting that Bell *et al*. (2020) were investigating cross-species data and population fluctuations within individual species may obscure any synchrony in trends. That being said, Estay *et al*. (2009) found spatial synchrony in oak aphid, *Tuberculatus annulatus* (Hartig) catches in UK suction-traps, declining with distance and approaching zero at around 360 km (comparable to Bell *et al*. (2020)'s 338 km). In contrast, Hanski & Woiwod (1993) using combined species data from the same networks of traps found spatial synchrony in both aphids and moths. Synchrony in aphids was stronger than in moths and both declined with distance but both taxa retained a low level of synchrony at over 600 km, the longest possible distance in the study region.

Other researchers have found spatial synchrony in insect populations using nonparametric covariance functions (ncf) to generate spline correlograms. Particular attention has been paid to outbreaking forest pest species (Williams & Liebhold 2000, Peltonen *et al*. 2002, Økland *et al*. 2005, Aukema *et al*. 2006, Choi *et al*. 2011, Lantschner *et al*. 2019) where measures of abundance were either qualitative, based on extent of damage (Williams & Liebhold 2000, Peltonen *et al*. 2002, Økland *et al*. 2005, Aukema *et al*. 2006, Lantschner *et al*. 2019) or quantitative, based on counts of pest-infested pine needles (Choi *et al*. 2011). All studies showed that synchrony declined with distance, but the rate and extent of this decline varied both between and within studies. Spruce budworm, *Choristoneura fumiferana* (Clemens), outbreaks were studied by Williams & Liebhold (2000) in North American forests where they found that synchrony decreased with distance, approaching randomness near 2000 km. Peltonen *et al*. (2002) found spatial synchrony in a range of outbreaking insect pests (mostly Lepidoptera, including the Williams & Liebhold (2000) dataset) in North America and Europe. Despite clear differences in dispersal ability between species the extent and level of synchrony did not vary significantly between them (Peltonen 2002). Økland *et al*. (2005) examined a suite of bark beetle species (*Dendroctonus* Erichson and *Ips* De Geer) in Europe and North America. They found that patterns of synchrony were very similar, however a significant difference between the species in terms of maximum level of covariance and the rate of decline with distance was reported (Økland *et al*. 2005). Aukema *et al*. (2006) found that regional synchrony in tree mortality caused by mountain pine beetle *Dendroctonus ponderosae* (Hopkins) attacks in British Columbia differed between incipient (non-outbreak) years and epidemic years, with synchrony becoming more widespread in years when the beetle outbreaks occurred (Aukema *et al*. 2006). Choi *et al*. (2011) used percentage infestation of the pine needle gall midge, *Thecodiplosis japonensis* Uchida & Inouye, to determine the spatial synchrony of the pest in Korea. They found that the extent of spatial synchrony and rate of decline differed significantly between regions (Choi *et al*. 2011). Lantschner *et al*. (2019) found a high level of spatial synchrony in the woodwasp, *Sirex noctilio* Fabricius, outbreaks in Patagonian pine plantations. Notably they found that, though synchrony shows some decrease with distance, it remained at a high level of covariance across the 300 km study region. Weather was implicated as a major driver of synchrony across these studies of outbreak forest pests, in particular temperature (Williams & Liebhold 2000, Peltonen *et al*. 2002, Aukema *et al*. 2006, Choi *et al*. 2011) and precipitation (Williams & Liebhold 2000, Peltonen *et al*. 2002, Choi *et al*. 2006, Lantschner *et al*. 2019), although Peltonen *et al*. (2006) found that synchrony dropped more rapidly in the taxa under consideration than weather variables, indicating other factors such as habitat quality were also important. In other systems, ncf splines have been used to show that synchrony in populations of the spotted-wing fruit fly *Drosophila suzukii* (Matsumura) declined with distance and altitude in all studied regions in Italy with the exception of the Dolomites where synchrony remained high at all altitudes, though this result should be treated with caution due to the short timescale of observations (Santoiemma *et al*. 2019). As described above, Bell *et al*. (2020) reported spatial synchrony in aphids in the UK, but not in moths using a 47-year dataset.

4.5.5 Implications for the use of suction-traps in a DSS

One of the aims of this study was to determine if suction-trap records could function as part of a DSS. If there is a relationship between suction-trap counts and numbers caught in the crop, then this suggests the suction-trap can be used to inform risk in the field. Additionally, if spatial synchrony existed in *B. aeneus* populations then at the very least the numbers in one field can predict the situation in a nearby field, thus only 1 field per farm needs to be sampled for decision-making. Indeed, it could be expected that if synchrony existed at the landscape level demonstrated for aphids in suction-traps (Hanski & Woiwod 1993, Estay *et al*. 2009, Bell *et al*. 2020), then one field or suction-trap can enable decision making at a regional level. However, I did not find evidence of spatial synchrony in *B. aeneus* at the field-scale, suggesting that decisions need to be made on a field-by field basis on each farm in each region. It could be hoped that, despite a lack of relationship between counts in the two networks studied here, crop inspections could still be informed by the findings in Chapters 2 and 3 of this thesis, in that the conditions that lead to higher counts, either seasonal or daily, provide a signal that crop inspections would be advised.

4.5.6 Next steps

The work in this Chapter could be continued with the aim of developing a further understanding of the spatial ecology of *B. aeneus*. The asynchronous sampling detailed in Fig. 4.2 may not have been adequately dealt with in the analysis and if a better weighting system for separating multi-day counts into standardised weeks could have been developed, there remains the possibility that synchrony at the field-level could have been identified. That being said, the method for assigning trap data to standard weeks – using the Central England Temperature series to remove colder days – would be expected to bring the field traps into a potentially spurious alignment that would artificially increase synchrony. Sticky-traps are a good measure of migration into a crop but may not reflect establishment within the field. Studies comparing aphid counts in suction-traps with field counts relied on in-field sampling of aphid infestations (Leather & Lehti 1982, McVean *et al*. 1999), though the most compelling relationship reported (Fabre *et al*. 2010) used alate migration into the crop as the field-level data, a system analogous to that used in this study. In-field transect data of *B. aeneus* on plants at the field sites in this study are available and it would be interesting to compare synchrony within those data, the dataset proved too difficult to analyse currently but this analysis remains an aspiration for future work.

Chapter 5.

Dispersal and migration of British pollen beetles

Chapter Summary

Tethered flight mills and molecular methods were used to assess the dispersal ability of Brassicogethes aeneus *and estimate their maximum flight distance.*

Brassicogethes aeneus *proved intractable for use on the flight mills and, as no data could be gathered, this study was discontinued with no results*

Individuals of B. aeneus *taken from suction-trap samples at five sites in 2018, and emergence trap samples from several fields at one farm site in 2017 and 2018 were investigated using next generation sequencing methods to determine population structure.*

Due to the poor quality of the degraded DNA extracted only a preliminary analysis could be completed. The results were inconclusive but low heterozygosity and high nucleotide diversity compared to published estimates of other coleopteran taxa indicates that the population of B. aeneus *in the UK is likely to be homogenous and that a proportion of beetles are able to migrate long distances.*

5.1 Introduction

5.1.1 Pollen beetle dispersal

Pollen beetles (*Brassicogethes (Meligethes) aeneus* (Fab.)) undergo two migration events every year, from and to overwintering sites (Williams *et al*. 2007, Williams & Cook 2010, Mauchline *et al*. 2013, Juhel *et al*. 2017, 2019). In addition, there is local dispersal to breeding habitat after a period of feeding on non-host plant pollen after emerging from hibernation and they may also disperse between and within the breeding habitat during the summer (Williams *et al*. 2007, Williams & Cook 2010, Mauchline *et al*. 2017).

5.1.1.1 Dispersal from overwintering grounds

Brassicogethes aeneus are thought to overwinter as adults in leaf litter or the top layer of soil (Hokkanen 1993, Alford *et al*. 2003, Rusch *et al*. 2012) of woodland and hedgerows (Fritszche 1957, Marczali 2006), or in winter oilseed rape (OSR) fields (Sutter *et al*. 2018). In early spring they emerge and migrate, either directly to the host brassicas including oilseed rape (Cook *et al*. 2004a, Williams *et al*. 2007, Williams & Cook 2010) or via pollen feeding on non-host plants in the wider landscape (Free & Williams 1978, Ouvrard *et al*. 2016). During the early spring there is a threshold temperature for flight (Mauchline *et al*. 2017), thought to be between 8°C and 15°C (Taimr *et al*., 1967; Tulisalo & Wuori, 1986; Láska & Kocourek, 1991; Šedivý & Kocourek, 1994, Junk *et al*. 2016). In the UK, migration into crops occurs around the end of April and early May (Mauchline *et al*. 2017), with reproduction beginning in June and July.

5.1.1.2 Movement between habitats in summer

Once the winter OSR has completed flowering, pollen beetles are known to move to spring OSR or native brassicas in July to continue breeding (Mauchline *et al*. 2017) or to other flowering plants in the landscape to continue feeding until their move to overwintering sites (Mauchline *et al*. 2017).

5.1.1.3 Dispersal to overwintering grounds

Dispersal to overwintering sites is poorly understood, both in terms of their location and the timing of the migration. What is understood is that they normally disappear from suction-trap catches in mid-August (this thesis Chapter 2) and are very rarely seen beyond the start of September.

5.1.2 What is known about the distance pollen beetles disperse

There have been several studies regarding the distance individual pollen beetles migrate with little consensus reached. Estimates range from 1.2 km (Juhel *et al*. 2017) to 13.5 km (Taimr *et al*. 1967) and Stechmann & Schütte (1976) reported them able to fly at least 12 km when released over the sea. There are indications that the dispersal range could be much greater: One beetle has been recorded in an offshore trap 80-160 km from land (Hardy & Cheng 1986) and individuals have been recorded in vertical-looking radar traces and high altitude (200m) nets (Chapman *et al*. 2004b, Mauchline 2003, Mauchline *et al*. 2017), with one caught at an altitude of 5000 feet (1524m) in an aeroplane-mounted aeroplankton sampling device (Weidel 2008). Even a short period of time spent at these heights would result in a passive horizontal movement of several kilometres (Johnson 1969, Reynolds *et al*. 2017).

5.1.3 Methods to study dispersal

There are a range of methods to study dispersal in invertebrates. The study of insect flight performance on flight mills (Rowley *et al*. 1968, Bordon & Bennett 1969, Jones *et al*. 2016, Minter *et al*. 2018) allows measurement of flight duration and speed in controlled laboratory conditions. Field studies on insect dispersal are more difficult, primarily due to their small size and large numbers (Osborne *et al*. 2002). Some of the main field methods are outlined below.

5.1.3.1 Observation

Tracking the movement of a study organism through direct observation is possible with invertebrates (Williams 1957, Schneider 1962, Southwood 1962, Wratten 2012), in practise it is difficult with all but the largest of invertebrates and can only be done on a relatively small scale (Osborne *et al*. 2002). A further impediment is the near impossibility of direct observation of insect dispersal at altitude (Johnson *et al*. 1962, Southwood 1962) unless the study insect is large and/or forms cohesive swarms, such as locusts (*Locusta migratoria* L.) for example (Rainey 1962).

5.1.3.1.1 Indirect Observation

Tracking invertebrates indirectly through trackable tags allows individuals to be followed over a greater distance (Riley 1995, Osborne *et al*. 2002, Chapman *et al*. 2004a). Radio tracking is rare in invertebrates (Hedin & Ranius 2002, Maitland *et al*. 2002, Godfrey & Bryant 2003, Brouwers *et al*. 2009). Tracking using harmonic radar (Mascanzoni & Wallin 1986, Riley *et al*. 1996, Lovei *et al*. 1997, Osborne *et al*. 1999, Colpitts & Boiteau 2004) is more common, however both methods are restricted by the size of the insect (Riley *et al*. 1995, Godfrey & Bryant 2003, Boiteau & Colpitts 2004), with pollen beetles too small to carry a tracker antenna at present, and restricted to tracking local movement rather than migration (Chapman *et al*. 2004a, 2011).

5.1.3.1.2 Radar

Radar aeroecology, the use of radar systems to observe airborne animals (Chilson *et al*. 2017) has become an important area of research in recent decades (Bauer *et al*. 2017). Tracking insects, either with weather radar (Nieminen *et al*. 2000, Drake & Reynolds 2012, Stepanian *et al*. 2016) or vertical-looking radar (VLR) (Chapman *et al*. 2011, Drake & Reynolds 2012), has been shown to be effective in forecasting the migrations of agricultural pests such as locusts (Drake & Wang 2013), aphids (Leskinen *et al*. 2011) and moths (Leskinen *et al*. 2011, Westbrook *et al*. 2014). Again, size is a restricting factor – the more so the greater the height at which the invertebrates fly (Chapman *et al*. 2011, Reynolds *et al*. 2017). In addition, the identification of the species responsible for a radar trace is difficult in most cases (Wood *et al*. 2009).

5.1.3.1.3 Modelling using point-sampled abundance.

Spatially explicit population models (Dunning *et al*. 1995) can use point-sampled abundance data, to combine population models with landscape-scale geographical data. These spatially explicit models can be used to explore the effects of landscape features on population dynamics, either through space or time (Dunning *et al*. 1995, Juhel *et al*. 2017). These models can also be used to infer unobserved processes, such as dispersal, based on comparisons between observed patterns in abundance (Barbu *et al*. 2010), and has been used to investigate pollen beetle dispersal in Normandy, France (Juhel *et al*. 2017) estimating mean dispersal at 1.2 km.

5.1.3.2 Mark-recapture and Capture-mark-recapture

As the perennial problem of insect size restricts our ability to directly observe insect movement other methods have been developed to infer the distance invertebrates migrate, or the potential distance they are able to migrate. The principal technique for many years has been to mark individuals or populations and record the distance they travel before recapture (Reynolds *et al*. 1997, Southwood & Henderson 2000, Bennetts *et al*. 2001, Hagler & Jackson 2001). Two classes of this method can be delineated: Mark-recapture methods involve tagging subjects in the field, usually without direct handling, whereas capture-mark-recapture (or mark-release-recapture) methods have a collection step, either from the laboratory or the field, to gather subjects to mark (Hagler & Jackson 2001). Mark-recapture methods include general location techniques using markers present in the environment and measured by examining gut contents (Reynolds *et al*. 1997); identifying pollen carried by the insect (Hendrix III *et al*. 1987); or analysis of the insect's elemental composition (chemoprint) to determine the likely provenance of captured specimens (Turner & Bowden 1983). This can be taken a step further by experimentally adding traceable material to the environment through diet supplements (Quarterman *et al*. 1954a, b, Jones *et al*. 1980) or as a substance such as glitter (Jackson *et al*. 1999), powder dye (Denholm *et al*. 1985, Byrne 1999, Southwood & Henderson 2000), radioactive markers (Quartermann *et al*. 1954a, b, Klick *et al*. 2014), fluorescent markers (Stechmann & Schütte 1976, Denholm *et al*. 1985, Jones *et al*.

1999); or stable isotopes (Graham *et al*. 1978) to be picked up by the insect and recovered post-dispersal. Capture-mark-recapture of insects can involve using tags (Urquhart & Urquhart 1979, Rubink 1988, Osborne & Williams 2001), directly marking individuals, sometimes called mutilation (Gangwere *et al*. 1964, Winder 2004, Griffiths *et al*. 2005), or applying internal markers through injection or feeding (Gangwere *et al*. 1964, Strand *et al*. 1990).

One of the drawbacks to these techniques is the low recapture rate, especially at greater distances (Hendrix III *et al*. 1987, Showers 1997, López-Pantoja *et al*. 2008). Attractant lures can increase capture rates but can introduce bias (Jones *et al*. 1999). Other problems include the expense of markers, both in researcher time and materials (Reynolds *et al*. 1997, Hagler & Jackson 2001), and the safety aspects of using radioactive or persistent materials (Reynolds *et al*. 1997). Additionally, it is still unclear whether individuals marked using capture-mark-recapture methods are affected behaviourally by the process (Morton 1982, Hagler & Jackson 2001).

5.1.3.3 Molecular methods

Because dispersal leads to geneflow when reproduction is successful it is possible to use genomics to infer dispersal rate or distance (Rousset 2001, Broquet & Petit 2009). As genes flow between populations, molecular markers can be used to assess allele frequencies and infer dispersal rate and/or distance (Wright 1943, Slatkin 1987, Broquet & Petit 2009). Measurements of genetic divergence between populations can be used to estimate levels of gene flow, and by extension infer dispersal capacity (Kim & Sappington 2013). Dispersal rate refers to two processes – forward and backward dispersal (Broquet & Petit 2009). Forward dispersal refers to the proportion of individuals emigrating from a population and is tricky, if not impossible to measure in insects (Broquet & Petit 2009) although it has been modelled using known dispersal distances in *Culicoides* midges (Eagles *et al*. 2014). The immigration into a population is known as backwards dispersal and is a more tractable measure across ecological systems and can be estimated using molecular markers (Broquet & Petit 2009).

There are several ways to estimate backwards dispersal from molecular markers (Broquet & Petit 2009). The most commonly used molecular markers include: microsatellite markers (Waser & Strobeck 1998), DNA fingerprinting, or barcoding, using markers such as AFLP (amplified fragment length polymorphisms) (Vos *et al*. 1995, Mueller & Wolfenbarger 1999) and CoI (cytochrome oxidase I) (Lunt *et al*. 1996, Hebert *et al*. 2003). Using these markers modelling approaches based on allelic identity and/or isolation by distance can be undertaken (Wright 1943, Cockerham 1973, Weir & Cockerham 1984, Slatkin 1987, Meirmans 2006). In pollen beetles, efforts have previously been made to estimate dispersal (Kazachkova *et al*. 2007, 2008, Makünas 2012, Ouvrard *et al*. 2016, Juhel *et al*. 2019), with mixed results (Table 5.1), though differentiation between different populations was generally low across all studies (Ouvrard *et al*. 2016).

Study authors	Markers used	Location	Distance estimate
Kazachkova *et al*. 2007	AFLP	Sweden	High local dispersal, but low gene flow between populations at a continental level.
Kazachkova *et al*. 2008	AFLP	Europe	
Makünas 2012	Mt DNA COI gene	Lithuania & Finland	
Ouvrard *et al*. 2016	AMOVA	Europe	Unclear, probably high with longitudinal, rather than latitudinal, migration a feature.
Juhel *et al*. 2019	Microsatellite markers	Europe	Continental scale.

Table 5.1: Summary of previous genetic research into *Brassicogethes aeneus* dispersal.

5.2 Aims and objectives

The main aim of this chapter is to understand the dispersal ability of *B. aeneus* in the UK and to attempt to elucidate their maximum flight distances.

The objectives were;

1) to use tethered flight mills to estimate maximum flight performance in terms of flight duration and speed

2) to use molecular methods (DNA sequencing) to estimate geneflow between populations within the UK

3) to use these data to infer dispersal distance of *B. aeneus*

5.3 Methods

5.3.1 Flight mill study

5.3.1.1 Sample collection

The RIS operates a 12.2m suction-trap (Macaulay *et al.* 1988) at Rothamsted for live sampling of aerial invertebrates. Samples were taken from this trap and swept from OSR and other vegetation around Rothamsted farm, Hertfordshire. Samples were taken in March and April of 2015 and 2016.

5.3.1.2 Flight mill

The flight mills used were adapted from Jones *et al*. (2016) (Patent: Lim *et al*. 2013) (Fig. 5.1) incorporating lighter weight metal than has been previously used. The mill consists of a pair of magnets between which is suspended the mill structure. The mill itself is comprised of a lightweight arm with attachment points at either end for an insect and a counterweight respectively, a central axis has a disc with a banded pattern so that it turns with the arm. In full operation, a light detector detects the movement of the disc to record the distance flown and the flight speed. On the day of sampling the beetles were chilled and attached to a spindle using superglue and then hung from the lightweight arm. Beetles were observed for 30 minutes to determine whether they would show intent to fly. Unfortunately, despite the lightweight materials employed and attempts made to alter temperature and light regimes around the mill to encourage flight, pollen beetles were either unable, or unwilling to fly when attached to the mill with only one beetle out of a total of 34 completing a single circuit of the mill. Therefore, no results on flight duration or speed could be gathered using this method.

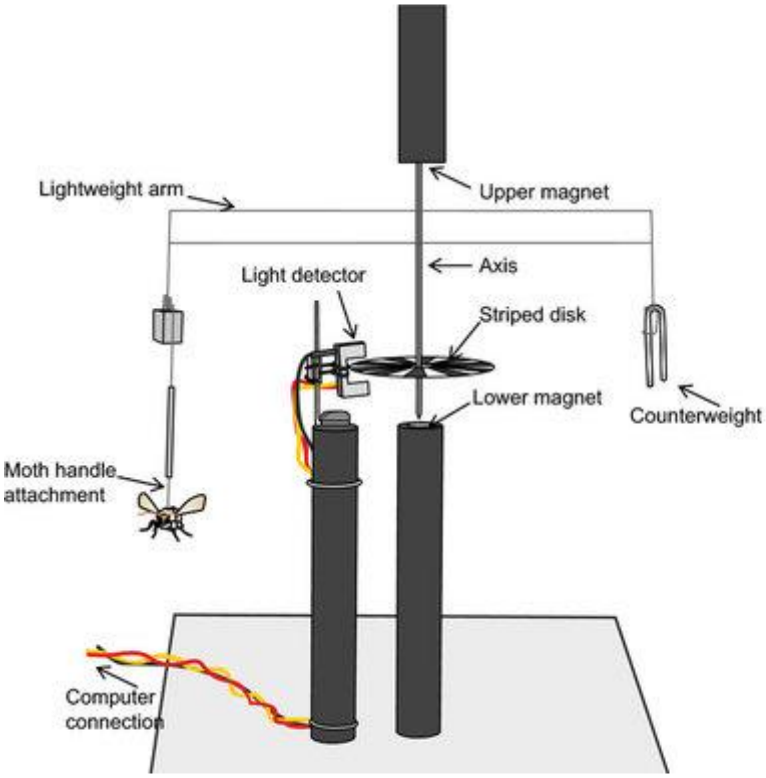


Fig. 5.1: Labelled diagram of flight mill design used for this study (Figure adapted from Jones *et al*. 2016).

5.3.2 Molecular study

5.3.2.1 Sample Collection

Beetles were taken from suction-trap and emergence trap (Fig. 5.2) samples in 2017 and 2018. The UK suction-trap network (Macaulay *et al*. 1988, Bell *et al*. 2015) operates traps in England and Scotland (Fig. 2.2). These traps (Fig. 5.2) are 12.2m tall and draw air at 45 m^3/min, capturing all insects that fly over the inlet pipe. Traps operate throughout the year, with samples changed daily during the peak aphid migration period (March-October) and weekly otherwise. Emergence traps were placed in oilseed rape (OSR) fields around the Rothamsted farm in 2017 and 2018 (Figs. 5.3 &

5.4). These traps capture insects emerging from eggs and pupae in the soil below them and were checked daily throughout the pollen beetle eclosion period (June-July). See Appendix C for more information on these traps. As the pollen beetle is a univoltine species that overwinters as an adult (Hokkanen 1993, Williams 2010), emergence trap samples from 2017 were paired with suction-trap samples from early 2018 (April-May) as “Generation 1” (those that emerged in 2017), with suction-trap samples from July 2018 paired with 2018 emergence trap samples as “Generation 2” (those that emerged in 2018). Suction-trap samples from five locations (Newcastle, York, Rothamsted, Wye and Starcross (See Figure 2.2)) were utilised, however a lack of successful DNA extractions from Starcross precluded using that site for “Generation 2”.

5.3.2.2 DNA extraction

DNA was extracted using the QIAamp® DNA Micro Kit [Qiagen, UK]. Prior to the extraction process, whole individual beetles were immersed for ~15 seconds in liquid nitrogen and homogenised using a sterile pestle. Homogenised samples were incubated at 56°C overnight in a solution of 180µl of ATL buffer with 20µl of Proteinase K.; 200µl of AL buffer with 1% RNA carrier was added to the sample then spun for 1 minute in a spin column at 8000 rpm. Respectively 500µl of AW1 buffer and 500µl of AW2 buffer were added to the sample and then spun down at 8000 rpm for 1 minute each time. Finally, samples were eluted in 30µl of AE buffer with a room temperature incubation for 10 minutes before being centrifuged at 14000rpm for 1 minute. The eluted samples were then incubated at room temperature for 10 minutes before being passed through the spin column a second time at 14000rpm for 1 minute to elute any DNA remaining in the spin column. Samples were treated with RNAase to remove carrier RNA, then stored at -80°C.

Fig. 5.2: Emergence trap (left) and suction-trap (right). Not to scale.

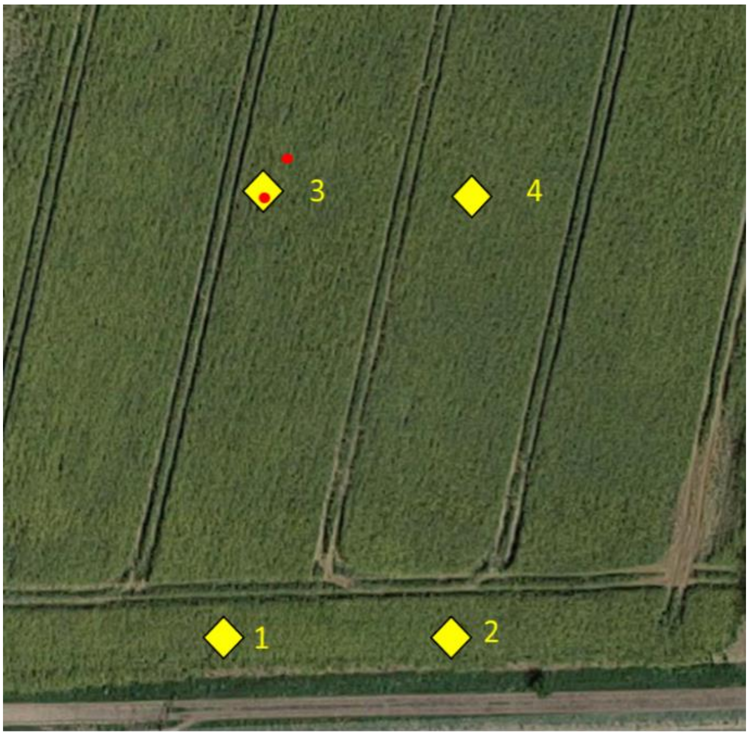


Fig. 5.3: Schematic of emergence trap locations within a field (yellow diamonds), temperature datalogger locations marked with red dots (see Appendix C for more information on the dataloggers).

Fig. 5.4: Map of Rothamsted farm with locations and names of oilseed rape field sites marked, boxed in yellow. Yellow star indicates the location of the Rothamsted Meteorological Station and the Rothamsted Insect Survey suction-trap compound.

5.3.2.3 Sequencing

Selected samples (Table 5.2) were genotyped using genome-wide single nucleotide polymorphisms (SNPs) which were identified using genotyping by sequencing (GBS), a genomic reduced representation sequencing method. Single digest restriction-site associated DNA (sdRAD) library preparation and sequencing were carried out by the Edinburgh Genomics facility at the

University of Edinburgh. DNA was digested with the six-base cutting enzyme EcoRI, which has a cut site approximately every 4kb pairs.

5.3.3.3 Analysis

The quality of the genomic sequences was investigated using FastQC. The resulting sequence dataset was analysed using a *de novo* approach with *Stacks 2.0* software (Catchen *et al*. 2011, 2013), starting with process_radtags to demultiplex the individual samples. Sequences were then analysed using the populations wrapper for the pipeline. In Stacks the sampled loci from each individual were reassembled *de novo* from the sequenced reads using *ustacks*. The loci assembled from each individual independently were then synthesised into a complete view (catalogue) of all loci and alleles present in the population using *cstacks* (Rochette & Catchen 2017). Using a population map text file containing two columns, one for the samples and the other for the population, populations estimates, genetic diversity statistics and divergence between populations were identified in the text file. There are further analytical steps to be completed on these data, however these will be done at a later date.

5.4 Results

In total, DNA was successfully extracted from 229 pollen beetles. Table 5.2 details the number of beetles from each site and generation, of these a subset was sent for radSeq (Table 5.2). The total number of beetles included in the genetic analysis was 54 (across both generations and 5 suction-trap sites); the analysis therefore looked at the diversity of the genetic makeup within this sample population.

DNA quality and quantity were lower than expected (Fig. 5.5) with evidence of molecular degradation in most samples. Due to difficulties encountered with the sequencing step as a result of the DNA degradation, currently only the Stacks step of the analysis has been completed and only the suction-trap samples have been analysed (Table 5.3). Given the degraded nature of the DNA, analyses have been carried out allowing five levels of mismatch in the catalogue assembly step in order to gain insight into the best level of sampling mismatches to accept in assembling the catalogue. Table 5.3 provides a summary of high-level parameters.

The number of SNPs ranged from 76 to 255 depending on how many alignment mismatches are allowed. This number reduces with increasing mismatch allowance as the degradation of nucleotides causes gaps within SNPs which leads to a lack of alignment that causes the Stacks software to classify what would be identical SNPs as different. The mean number of samples per locus remains stable at between 44 and 45. The variant position (Pi) – a measure of nucleotide diversity (Nei & Li 1979) is between 0.02903 and 0.0355. Expected Heterozygosity (H_E) is 0.0287-0.0351 and observed Heterozygosity (H_O) 0.0268-0.02968. These heterozygosity measures are used to assess genetic variation within a population (Harris & DeGiorgio 2017). Low values mean low diversity and can be a sign of recent bottlenecks in population size (Chakraborty & Nei 1977, Allendorf 1986). The comparison of H_O and H_E gives insight into the nature of the population under study. If H_O is lower than H_E this may indicate high levels of inbreeding within a population (Chakraborty 1987, Templeton & Read 1994) or clonal reproduction, such as that found in aphids (Morales-Hojas *et al*. 2020). Here the values are similar, suggesting there is little inbreeding occurring in the population. Similarly, a high value for the inbreeding coefficient (Fis) can indicate lower genetic diversity than would be expected in a normal population (Charlesworth 2003), here Fis ranged from 0.02276 to 0.18201 depending on the number of mismatched loci allowed. Again, these values are low and indicate that inbreeding is rare in UK *B. aeneus.*

Site	Generation 1	Generation 2
Rothamsted field sites *(not included in the Stacks analysis)*	32 (18)	48 (28)
Rothamsted suction-trap	20 (7)	12 (7)
Starcross suction-trap	10 (6)	0 (0)
Wye suction-trap	20 (5)	10 (7)
York suction-trap	30 (5)	10 (6)
Newcastle suction-trap	25 (6)	12 (5)
Total number of beetles included in the Stacks analysis	29	25

Table 5.2: Number of beetles sampled from each location and each generation and (in parentheses) number sequenced.

No. of mismatches allowed between sample loci	**1**	**2**	**3**	**4**	**5**
No. genotyped loci	902,810	744,846	661,961	612,105	582,879
Mean coverage loci	8.1x	8.1x	8.2x	8.2x	8.1x
No. polymorphic loci shared between 80% samples in the population	121	79	59	31	31
No. SNPs (variant sites)	255	177	116	81	76
Mean samples per locus	44.631	44.621	44.879	44.556	44.697
Pi (variant positions)	0.02903	0.032	0.03023	0.0313	0.0355
Expected Heterozygosity (H_E) **(variant positions)**	0.0287	0.03164	0.02989	0.03095	0.0351
Observed Heterozygosity (H_O) **(variant positions)**	0.02682	0.02725	0.02482	0.02968	0.0268
Fis (variant positions)	0.04616	0.06225	0.09996	0.02276	0.18201

Table 5.3: Summary statistics of genetic diversity in the suction-trap samples as measured by Stacks. No. of mismatches refers to the number of incorrect matches allowed (given the degraded nature of the DNA this leads to more alignment and a reduced number of loci). No. genotyped loci is the sum total of all loci identified; these have a mean coverage of ~8. The number of polymorphic loci shared between 80% of samples is between 31 and 121, these numbers drop as the number of mismatches increases.

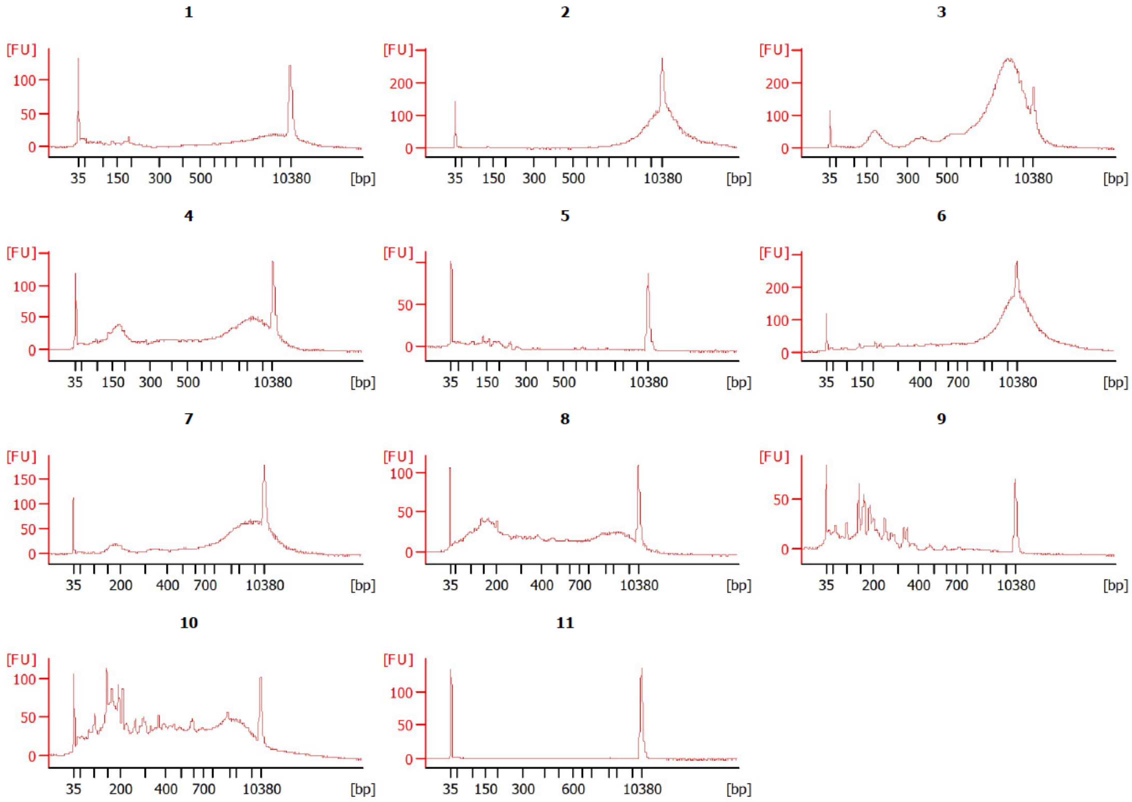


Fig. 5.5: High sensitivity DNA assay for 10 samples from Table 5.2. The x-axis shows the size of the DNA fragments between 35 and 10380 base pairs (bp), the y-axis shows the sample intensity (FU- fluorescent units), a measure of the amount of DNA at each size in the sample. Samples 5 and 9 have no large size DNA and appear highly degraded; Samples 2 and 6 show high molecular weight DNA with small amounts of degradation; Samples 1,3,4,7,8,10 show varying levels of degradation; Sample 11 is the control/reference.

5.5 Discussion

In the absence of a robust and comprehensive analysis of the degraded DNA, and with the failure of the flight mills to provide data at all, conclusions drawn from this research must be considered tentative.

The failure of the flight mills is particularly disappointing as this method has yielded robust results for flight distance in other beetle taxa (e.g. David *et al.* 2014, Kees *et al.* 2017, Barkan *et al.* 2019), albeit the species surveyed are much larger in size than *B. aeneus.* It was clear that, although methods were made to encourage flight in the lab, for example altering time of day, changing the light levels and adding flowers of OSR to the vicinity of the mills, the beetles were unable to fly. Whether this inability was behavioural or physical remains unclear.

The similarity between the H_O and H_E values indicates that this is likely a normal sexual population, or populations, in equilibrium in that there is no evidence for lower, or greater levels of genetic diversity than should be expected. These values appear to be low when compared to published estimates of heterozygosity in other beetle taxa. For example Desender *et al*. (1988) reported H_E of 0.1 and 0.3 for two species of saltmarsh carabid beetle; Desender *et al.* (2005) reported H_E of 0.126-0.256 and 0.013-0.115 for two species of flightless carabid beetles; Grapputo (2005) reported that Colorado beetles (*Leptinotarsa decemlineata* Say) had H_E of 0.192-0.246 in their native range and 0.013-0.115 in their invasive range; Rafter *et al.* (2018) reported H_E of 0.662-0.679 and 0.858-0.887 in two species of grain beetle pest.

Conversely the values for nucleotide diversity (Pi) reported here are higher than many other beetle taxa in the literature. For example Cai *et al*. (2008) reported Pi at levels of 0.003-0.014 in red turpentine beetles (*Dendroctonus valens* LeConte) in their native and invaded ranges; Giska (2015) reported Pi as 0.00075-0.003 in staphylinid beetles; Crossley *et al*. (2019) reported Pi to range between 0.0056 and 0.008 in the Colorado beetle.

The next steps in the analysis of these genetic data are firstly to filter the single-nucleotide polymorphisms (SNPs) using vcftools v0.1.14 (Danecek *et al*. 2011) to maximise the quality of the SNPs and genotypes, whilst minimising missing data at the individual, or marker level as recommended by O'Leary *et al*. (2018). To improve the chance of detecting subtle genetic structure the sampling locations will be set as priors, allowing the detection of a weak structure without biasing the results towards detecting structure when there is none (Morales-Hojas *et al*. 2020). The filtered SNPs can then be analysed using the Structure 2.3.4 software system (Pritchard *et al*. 2000), a Bayesian genetic clustering algorithm, to estimate geneflow between *B. aeneus* populations within the UK.

The aim of this chapter was to determine the population structure of the British *B. aeneus* populations and, from that, to attempt to get an estimate of the amount and distance of dispersal in the species. Unfortunately, due to the preliminary nature of the analysis thus far undertaken these aims are not yet met, however the initial data suggest that there is a single population of *B. aeneus* in England that is sufficiently mobile between the sample locations to allow genetic mixing. There is not yet sufficient evidence that the English *B. aeneus* population has a different structure to the European population research previously published (Table 5.1), but future analysis of these data samples should allow investigation into the geneflow between the sample location and therefore provide an estimate of the dispersal ability of this species within the UK.

Chapter 6.

General Discussion

6.1 Primary findings and contributions of this study

This study aimed to gain an improved understanding of the temporal and spatial population dynamics of *Brassicogethes aeneus* (Fab.), a pernicious pest of oilseed rape (OSR; *Brassica napus* L.). The ultimate aim was to use this knowledge to improve the management of *B. aeneus* in OSR and thereby lead to improved performance of farms through reduced insecticide inputs. I approached this by gathering novel data on key population parameters, modelling the population dynamics of this species in the UK, developing predictive models that can inform decision support tools enabling farmers to modify their management practices. Specifically, I wanted to explore the potential of using suction-traps to investigate the drivers of differences in abundance in *B. aeneus* populations at both annual and daily timescales in order to improve the quality and range of decision support systems (DSSs) available for control of this pest in the UK. To do this I used suction-trap data from the longest-running terrestrial insect dataset in the world (Bell *et al*. 2020) alongside field-collected data from an extensive sampling campaign comprising 178 sites (Skellern *et al*. 2017). In addition, I used cutting-edge techniques in flight-mills (Jones *et al*. 2016) and next-generation sequencing (Miller *et al*. 2007) to examine the flight behaviour and migratory potential of this species. Using a series of modelling approaches, my research has furthered the understanding of *B. aeneus* ecology and I summarise here the impact this will have on achieving my aim of improved pest management.

In Chapter 2 I established, from analysis of suction-trap data, that *B. aeneus* populations have been increasing in the UK along with an increase in cropped area of OSR. I could not demonstrate a link between these two trends, perhaps because *B. aeneus* populations continue to increase beyond 2011 when the trend for OSR reaches asymptote. This may indicate that *B. aeneus* populations may not yet have reached equilibrium with habitat area in the UK (i.e. the carrying capacity, or maximum population size that can be supported sustainably within the ecosystem, for *B. aeneus* in new OSR habitat has not been filled (Hixon 2008)), alternatively it may be the recent development of pyrethroid insecticide resistance in *B. aeneus* (Zimmer & Nauen 2011a) that has led to increased populations. Either way, this apparent continued increase in *B. aeneus* may have severe consequences for growers as this could mean increased yield loss as larger populations overwhelm permitted control measures (Hokkanen 2000, Zlof 2008) and a larger pool of individuals allowing greater adaptation within the population (Neher *et al.* 2010, Neher & Shraiman 2011) and thus the increased likelihood of strains resistant to insecticides developing.

Importantly for the development of DSSs for growers, I demonstrated that the abundance of *B. aeneus* migrating in the spring – the time when infestation in crops could cause yield loss – could be largely explained by the number of beetles caught in the suction-traps the previous summer. This suggests that there is potential to estimate the abundance of the spring migration in the autumn of the previous year. If this potential is realised in a DSS, either through the adaption of current models (e.g. the proPlant DSS now on Xarvio (Johnen *et al*. 2010)), or through development of a new model this would allow growers to make decisions based on earlier knowledge than previously available. Reports on *B. aeneus* abundance released in late summer could allow growers to adjust their rotation plans accordingly to choose an alternative to OSR in years of high forecast abundance or to increase OSR planting when forecast abundance is low. This could also allow growers more flexibility when it comes to whether to order insecticide in advance or save money by not purchasing product. It would also potentially allow growers to be ready and prepared to deal with *B. aeneus* problems in a timely manner. In the UK there are no pest forecast models able to predict pest pressure on an as yet unsown crop, although in certain systems research is ongoing with some success, for example in wheat bulb fly (*Delia coarctata* Fállen) (Leybourne *et al.* 2021). Elsewhere an autumn forecast model is available for sugarbeet weevil (*Bothynoderus punctiventris* Germ.) in the Ukraine, although details are scant (Fedorenko 2006, FAO 2021).

There is a clear need to understand more fully the daily dynamics of pest migration, particularly days of mass migration – migration events where total numbers in flight are measured in orders of magnitude over normal levels of abundance. Such mass migrations can lead to serious damage to crops (Markkula *et al.* 2008) and cause field populations to rapidly pass threshold levels for control in crops (Mumford & Norton 1984, Pedigo *et al.* 1986). Current DSS forecasting tools for *B. aeneus* can forecast the phenology of migration to a reasonable degree of accuracy (Collier *et al.* 2001, Johnen *et al.* 2010) but are not designed to give an explicit daily forecast of when a mass migration event may occur (Ferguson *et al.* 2013). Other DSSs based around observations, for example the Xarvio water trap system ((https://www.xarvio.com/gb/en/products/scouting/yellow-trap-analysis.html), or the Rothamsted Insect Survey (Harrington & Woiwod 2007) can give alerts pertaining to mass migration but, by definition, these are after the event. To attempt to gain a better understanding of daily *B. aeneus* migration dynamics to improve forecasting model in Chapter 3 I modelled the daily abundance of *B. aeneus* in suction-trap samples. This proved a challenging dataset characterised by large numbers of zeroes with sporadic peaks of very high counts and regular day-to-day changes of two or more orders of magnitude in count numbers. Despite these challenges, through various analytical approaches that attempted to account for the stochasticity and zero-inflated nature of the data, I found evidence that daily trap counts were driven by a small range of environmental variables (sunshine, wind speed and minimum or mean temperature), alongside the time of year (day number) and the number of beetles caught after June 1st the previous year – this last finding reinforces the conclusion from Chapter 2 that assessing the abundance in the suction-traps of the newly emerged generation in the late summer may prove important in predicting the abundance of the spring migration. However, none of the models could predict the very high counts that constitute a mass migration event in the suction-traps. These events are rare and suggest that, whilst it may be possible to set out the conditions that favour mass migration events in *B. aeneus,* actually forecasting when these would occur may be less tractable using the frequentist statistical modelling and machine learning approaches considered here. An analogous problem that continues to vex the best scientists is volcanic eruptions, which despite research being highly instrumented and well-funded, has yet to deliver a reliable forecasting system (Palmer 2020). Forecasting rare events that seemingly do not have any precursors (i.e. in this instance zero one day, several hundred the next) will remain a challenge for science and may only be furthered with new real-time monitoring field data at a high resolution (i.e. hour-by-hour) to detect the upswing in movement.

Chapter 4 focussed on determining whether suction-trap data could be related to field-sampled data and at what distance. I found that there was generally a strong relationship between suction-trap and field data at the weekly scale even at distances of more than 100 km. In addition, the analysis of the suction-trap data found that some spatial synchrony of migrating beetles was evident up to ~150 km.

However, there was a great deal of scatter in the relationship between the suction-trap counts and field counts and this relationship also varied from year to year reducing confidence in the result. Closer inspection of the spatial dynamics of *B. aeneus,* as measured by field-caught beetles, revealed that they appeared to have no spatial synchrony at the field level, suggesting that it is not possible to forecast population abundance in a field, even if using data sampled as close, geographically, as a neighbouring field. In the suction-traps some spatial synchrony of migrating beetles was evident up to ~150 km. If beetles are as locally irruptive and field outbreaks are effectively independent as these results suggest, this means that control decisions by the grower must be made on a field-by-field basis on each farm in each region rather than be driven by a generic, national DSS based on suction-trap counts. However, the findings in Chapters 2 & 3 identified the conditions (based on both meteorological indicators and suction-trap abundance data from the previous year) that lead to higher abundance of *B. aeneus* and spring migration events. Therefore, models based on suction-trap data and localised meteorological data could potentially provide farmers with a customised warning signal to begin crop inspections.

The dispersal ability of *B. aeneus* was investigated in Chapter 5 using flight-mill and next-generation sequencing techniques. *Brassicogethes aeneus* proved unwilling or unable to fly on flight mills with only a single full circuit performed across more than 30 test flights. In addition, an analysis of the molecular data from the next generation sequencing approach could not resolve the problem of the degraded DNA and the biostatistical pipeline that would inform levels of population mixing remains incomplete. Yet from the preliminary analysis of the diversity in the genetic data, in comparison to published estimates of other beetle taxa, heterozygosity appears to be low, whilst nucleotide diversity appears to be high indicating that it is likely there is a single UK population and suggesting that a proportion of the population migrates long distances within the UK. It should be stressed that these are preliminary findings and it would be unwise to draw any further conclusions from these results.

Appendix D presents a suite of *B. aeneus* lifecycle parameters that could, in future, inform a process-based model of *B. aeneus* that, in turn, could provide the basis for a DSS for OSR and other brassicaceous crops much like that used in the Warwick MORPH model (Collier 2001). These parameters cover the entirety of the beetle's lifecycle, the egg and larval stages have particularly detailed parameters gleaned from field and laboratory experiments and most authors are in general agreement. Understanding of the pupal and adult life stages is less comprehensive with some aspects only found in single-paper studies and there is a need for further work on the understanding of the adult stages in particular.

Together the results presented in this thesis enhance our fundamental knowledge of the ecology of this pest species and can be used to improve management decision making in the future.

6.2 Further work

In addition to our enhanced knowledge of the ecology of *B. aeneus,* this study has made one major finding that is of great importance for farmers' management of this pest species; the number of *B. aeneus* in spring migrations can be qualitatively forecast by the counts of beetles made in the summer the previous year, modulated by a suite of environmental variables that are relatively straightforward to collect. There is potential here to develop an early warning system for growers; either a qualitative forecast, based on suction-trap counts alone released at the end of summer for the following year, or a more detailed forecast released in early spring, based on trap numbers and weather forecasts. Further research is needed into the desirability of these forecasts, both in terms of usefulness and practicality.

This study could not conclusively demonstrate a link between counts of migrating beetles in suction-traps and subsequent local field infestations, in part due to the large variability and sampling issues in the field data, and this should be a prime focus of any further work. A field campaign collecting daily samples, or perhaps using new pest detection technologies such as LIDAR (Jansson & Brydegaard 2018, Hassell *et al.* 2021, Kirkeby *et al.* 2021) that would cut down on identification and processing time, could make this link explicit and enable the closer investigation of the apparently chaotic nature of field-caught *B. aeneus* data.

Further work on the molecular aspects of Chapter 5 would allow a full analysis of the genetic structure of *B. aeneus* in the UK. As suggested from the preliminary results, it is unlikely that these further analyses will reveal anything that differentiates the UK population from the European populations in previous research (Kazachkova *et al*. 2008, Ouvrard *et al*. 2016, Juhel *et al*. 2019), however if there is demonstrated to be multiple genetically distinct populations in the UK this could have implications for control, as currently control strategies are based on the assumption that the European population is relatively homogenous.

The currently available modelling approaches used in Chapter 3 were unable to forecast rare, extreme population events – especially at traps where they had not previously occurred. This has important implications for the population modelling of any pest species where the population is increasing; as underestimation of pest pressure in a DSS could lead to a lack of control efficacy in the field and/or an erosion of trust in the DSS to the point that it has no use.

This work has shown the potential for suction-trap data to provide unique data sets for investigation into the temporal and spatial ecology of *B. aeneus.* Further exploration and analysis of suction-trap collected data may yield yet more ecological and management perspectives. For example, it may be possible to investigate suction-trap samples from warmer regions, or regions with a higher proportion of spring OSR, to allow a better understanding of *B. aeneus* populations in future UK environmental change scenarios. In addition, further investigation of the suction-trap samples may provide insight into the population dynamics of an increasing pest species. Rothamsted Insect Survey has stored samples dating back to 1974 (Shortall 2013) and these could be investigated to determine how *B. aeneus* populations developed as OSR became established as an important crop in the UK. This in turn may inform the protection of novel crops from pests in the future.

Appendices

Appendix A, Chapter 2

This chapter explored the the national and regional trends in oilseed rape cultivation in the UK and long-term pollen beetle data for two sites; Rothamsted and Broom's Barn.

Libraries

Here are the R libraries required for this chapter:

```
library("poptrend")

library("mgcv")

library("tree")

library("fitdistrplus")

library("tidyr")

library("dplyr")

library("itsadug")

library("ggpubr")

library("gridExtra")

library("grid")
library("ggplot2")
library("lattice")

library("ggplotify")
```

The "tidyr" and "dplyr" libraries are used to rearrange a dataset if necessary. The "fitdistrplus" library is used to assess the statistical distribution. The "poptrend" and "mgcv" libraries are used to perform the analysis, the rest are for producing the tabular and graphical output.

Data

The following data files will be used. These data are available on request.

```
setwd("P:/Final datasets")
PBdist2<-read.csv("Rothamsted.csv")
OSR<-read.csv("OSR and PB data 1987 to 2019.csv")
AOSR<-read.csv("OSR data England.csv")
OSR2<-read.csv("OSR data England 2.csv")
OSR3<-read.csv("OSR Regions.csv")
OSR4<-read.csv("OSR and PB data 2001 to 2015 East.csv")
OSR5<-read.csv("2015 total with regional sown area.csv")
PB<-read.csv("All sites spring & summer & overwinter.csv")
TA3<-read.csv("Annual RT BB with NA new.csv")
SPRING<-read.csv("LSG long term data.csv")
Combined<-read.csv("Combined.csv")
```

Fitting statistical distribution

First, the statistical distribution of the data is assessed (Note, the data used to fit distributions here (PBdist) is an example; this step of assessing the statistical distribution was done for all datasets but for brevity will only be demonstrated here once.)

```
attach(PBdist2)
```

It is necessary to remove "no trapping" records (here classed as "NA"s)

```
PBdist2NoNA<-PBdist2%>% drop_na()
summary (PBdist2NoNA)

##   Julian.Date        X2015             X2016            X2017
##  Min.   : 40.0   Min.   :  0.000   Min.   : 0.000   Min.   : 0.000
##  1st Qu.:108.5   1st Qu.:  0.000   1st Qu.: 0.000   1st Qu.: 0.000
##  Median :177.0   Median :  0.000   Median : 0.000   Median : 0.000
##  Mean   :177.0   Mean   :  5.411   Mean   : 4.178   Mean   : 3.967
##  3rd Qu.:245.5   3rd Qu.:  1.000   3rd Qu.: 0.000   3rd Qu.: 2.000
##  Max.   :314.0   Max.   :109.000   Max.   :86.000   Max.   :85.000
```

This dataset had too many zeroes for a statistical distribution to be resolved. It underwent two further filtering steps to obtain a workable dataset:

```
PB20171stcut<-select(filter(PBdist2NoNA, Julian.Date >67), c(Julian.Date,
X2017))
PBSpring<-select(filter(PB20171stcut, Julian.Date<154), c(Julian.Date, X2017))
```

Then we attempt to fit a range of statistical distributions to the data (note, for brevity these are set not to run in markdown)

```
#GammaPBcountSpring <- fitdist(PBSpring$X2017, "gamma", discrete = FALSE, lower
= c(0,0), start = list(scale = 1, shape = 1))
#WeibullPBcountSpring <- fitdist(PBSpring$X2017, "weibull", discrete = FALSE,
lower = c(0,0), start = list(scale = 1, shape = 1))
#LognormalPBcount <- fitdist(PBSpring$X2017, "lnorm")
#normPBcount <- fitdist(PBdist$X2017, "norm", discrete = FALSE, lower = c(0,0))
#expPBcount <- fitdist(PBSpring$X2017, "exp", discrete = FALSE, lower = c(0,0))
#poisPBcount <- fitdist(PBSpring$X2017, "pois", discrete = TRUE)
#cauchyPBcount <- fitdist(PBSpring$X2017, "cauchy", discrete = FALSE, lower =
c(0,0), start = list(scale = 1, shape = 1))
#logisPBcount <- fitdist(PBdist$X2017, "logis", discrete = FALSE, lower =
c(0,0), start = list(scale = 1, shape = 1))
#nbinomPBcount <- fitdist(PBSpring$X2017, "nbinom", discrete = FALSE, lower =
c(0,0), start = list(scale = 1, shape = 1))
#geomPBcount <- fitdist(PBSpring$X2017, "geom", discrete = FALSE, lower =
c(0,0), start = list(scale = 1, shape = 1))
#betaPBcount <- fitdist(PBSpring$X2017, "beta", discrete = FALSE, lower =
c(0,0), start = list(scale = 1, shape = 1))
#invgammaPBcount <- fitdist(PBdist$X2017, "invgamma", discrete = FALSE, lower =
c(0,0), start = list(scale = 1, shape = 1))
#llogisPBcount <- fitdist(PBSpring$X2017, "llogis", discrete = FALSE, lower =
c(0,0), start = list(scale = 1, shape = 1))
#invweibullPBcount <- fitdist(PBSpring$X2017, "invweibull", discrete = FALSE,
```

```
lower = c(0,0), start = list(scale = 1, shape = 1))
#paretolPBcount <- fitdist(PBSpring$X2017, "paretol", discrete = FALSE, lower =
c(0,0), start = list(scale = 1, shape = 1))
#paretoBcount <- fitdist(PBdist$X2017, "pareto", discrete = FALSE, lower =
c(0,0), start = list(scale = 1, shape = 1))
#lgammaPBcount <- fitdist(PBSpring$X2017, "lgamma", discrete = FALSE, lower =
c(0,0), start = list(scale = 1, shape = 1))
#trgammaPBcount <- fitdist(PBSpring$X2017, "trgamma", discrete = FALSE, lower =
c(0,0), start = list(scale = 1, shape = 1))
#invtrgammaPBcount <- fitdist(PBdist$X2017, "invrtgamma", discrete = FALSE,
lower = c(0,0), start = list(scale = 1, shape = 1))
```

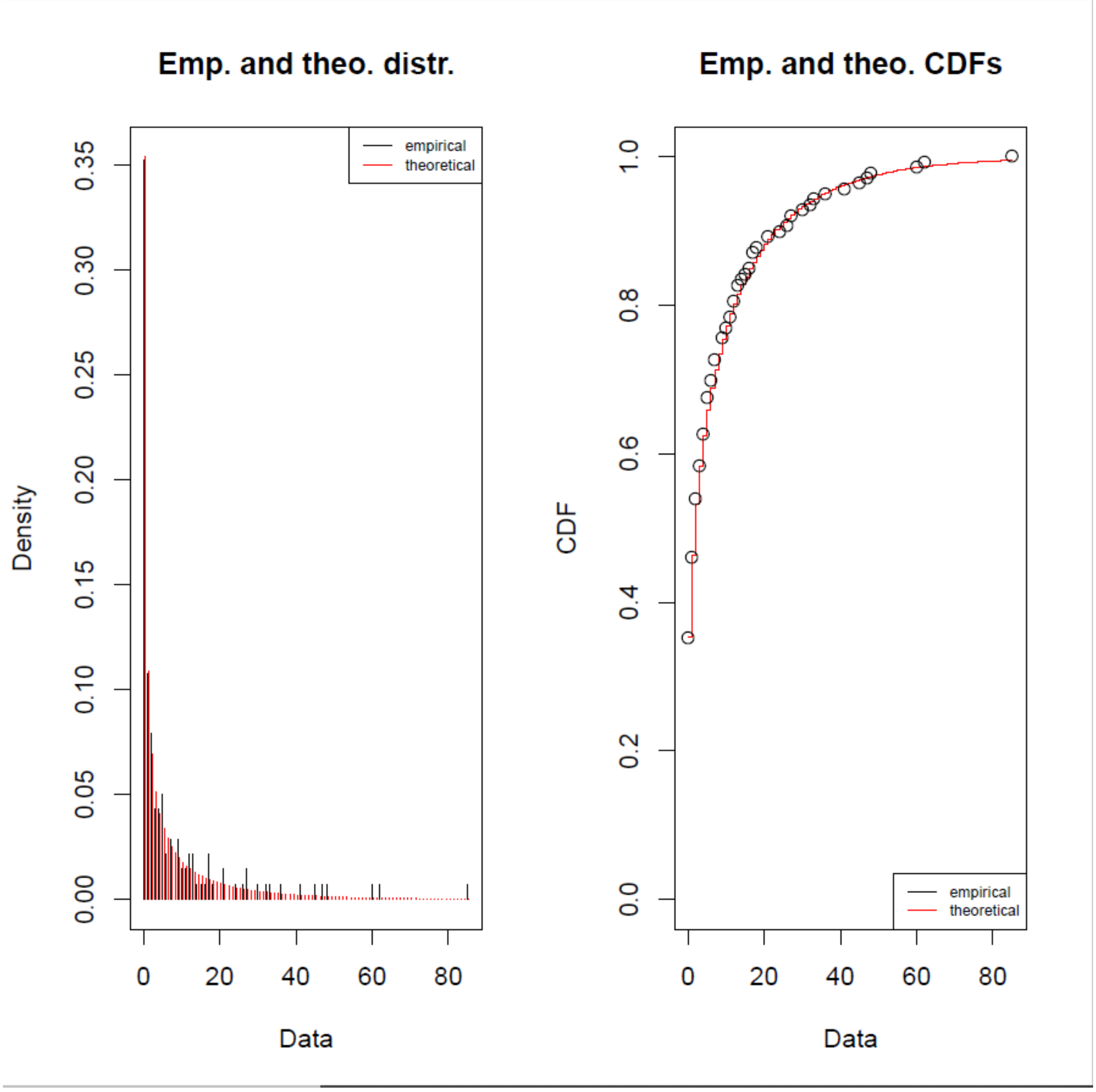
Then we produce graphical output for each successfully completed test (again, here for illustration purposes only for brevity):

```
#plot(GammaPBcountSpring)
#plot(WeibullPBcountSpring)
#plot(LognormalPBcount)
#plot(normPBcount)
#plot(expPBcount)
#plot(poisPBcount)
#plot(cauchyPBcount)
#plot(logisPBcount)
#plot(nbinomPBcount)
#plot(geomPBcount)
#plot(betaPBcount)
#plot(invgammaPBcount)
#plot(llogisPBcount)
#plot(invweibullPBcount)
#plot(paretolPBcount)
#plot(paretoPBcount)
#plot(lgammaPBcount)
#plot(trgammaPBcount)
#plot(invrtgammaPBcount)
```

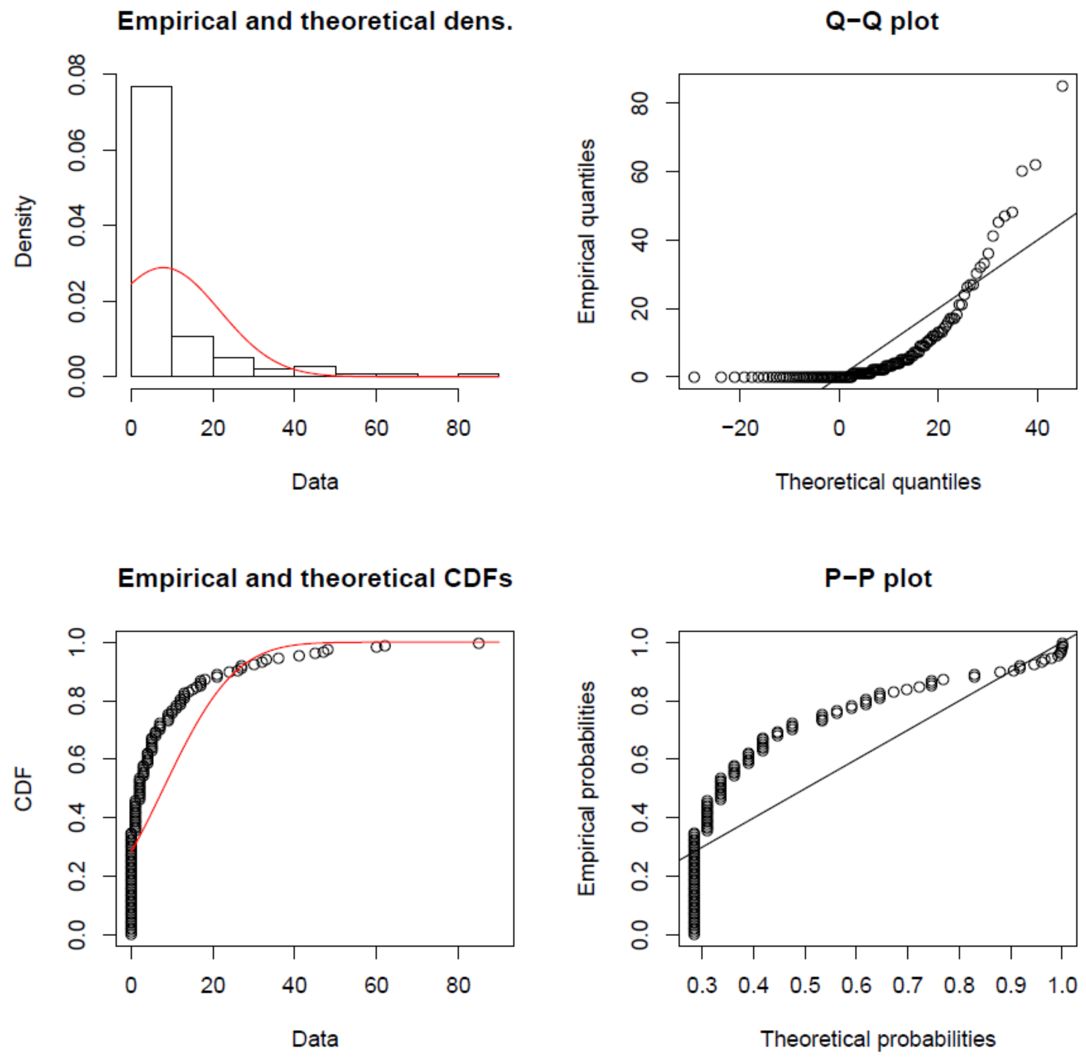
(Note, all graphical output is saved as a PDF file like the example code below. This is done for all thesis plots, but the code will not be presented again for the sake of reducing repetition).

```
#pdf(file = "P:/Analysis/fitdistgraphs/NegBinomial.pdf")
#plot(nbinomPBcount)
#dev.off()
```

Example graph:



This graph shows that the data follow a negative binomial distribution, compare with the following that attempts to fit a normal distribution to the data and fails:



Oilseed rape trends over time

Section 2.4.2

Here, the poptrend package is used to investigate the trend in cultivation of oilseed rape over time.

First some standard plots to give an overview of the data.

Fig. 2.5:

```r
plot(OSR$Year, OSR$Area, type = "l", col = "black", bty = "n", ylab = "Area of
OSR sown, '000 hectares")
```

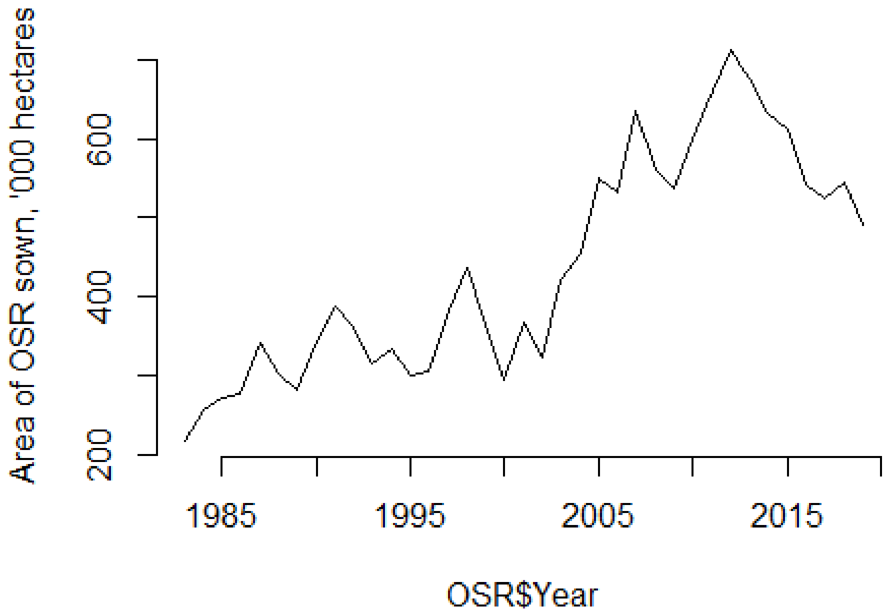
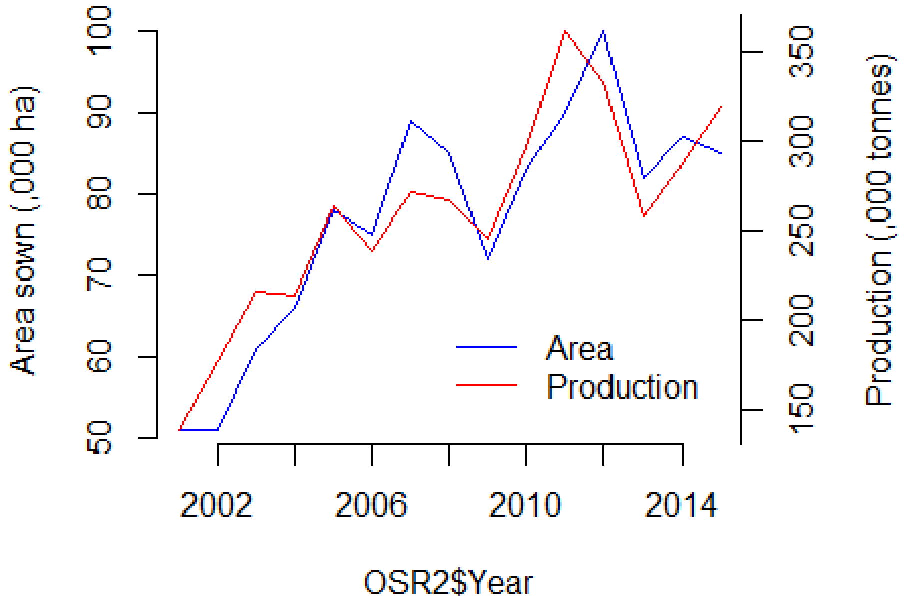


Fig. 2.6:

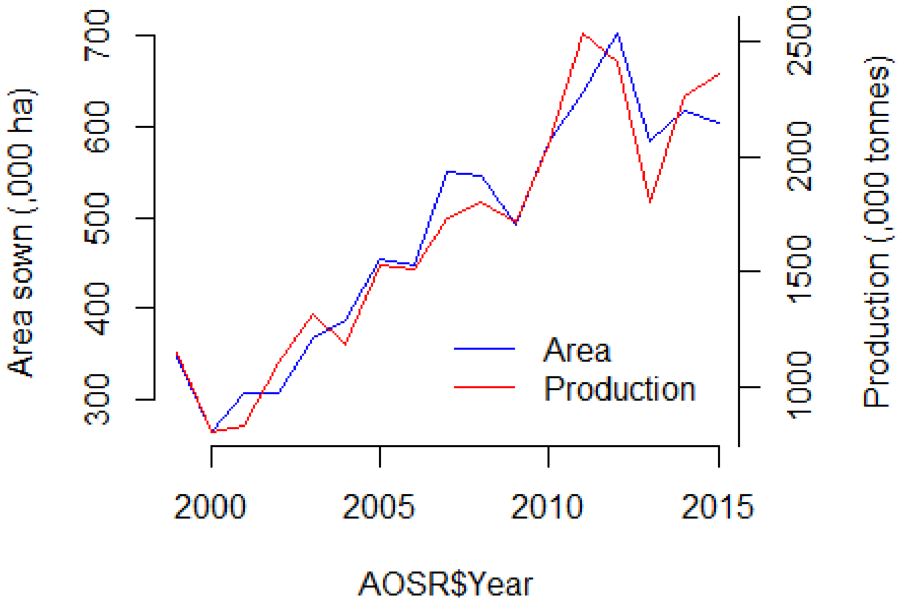
```
par(mar=c(5, 4, 4, 6))
plot(OSR2$Year, OSR2$Area, type = "l", col="blue", bty="n", ylab="Area sown (,000 ha)")
par(new=TRUE)
plot(OSR2$Year, OSR2$Production, type="l", axes=FALSE, bty="n", xlab = "", ylab = "", col="red")
par(new=FALSE)
axis(side=4, at=pretty(range(OSR2$Production)))
mtext("Production (,000 tonnes)", side=4, line=3)
legend("bottomright", inset=.05, cex=1, bty="n", c("Area", "Production"), lty=c(1, 1), lwd=c(1,1),col=c("blue", "red"))
```



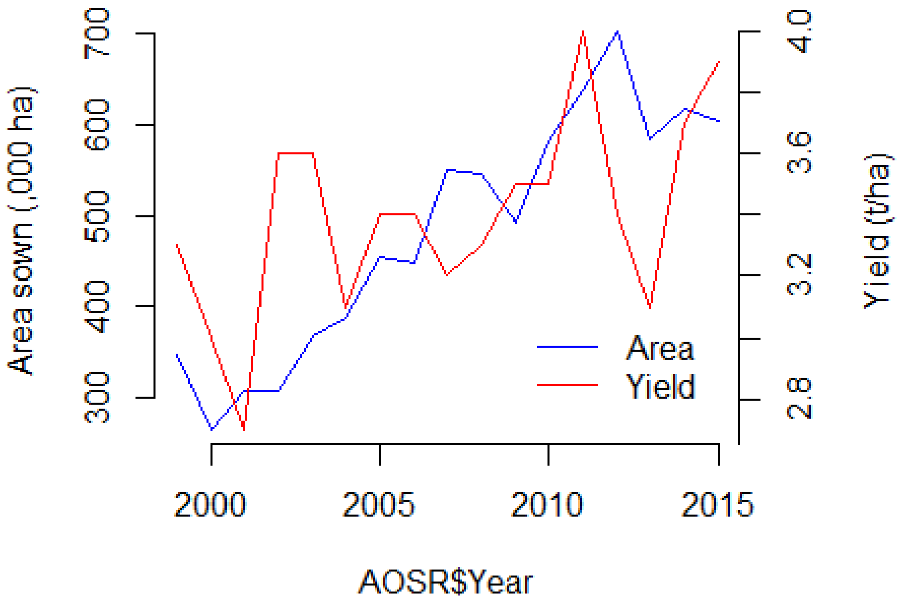
```
par(mar=c(5, 4, 4, 2))
```

Fig. 2.7:

```
par(mar=c(5, 4, 4, 6))
plot(AOSR$Year, AOSR$Area, type = "l", col="blue", bty="n", ylab="Area sown
(,000 ha)")
par(new=TRUE)
plot(AOSR$Year, AOSR$Production, type="l", axes=FALSE, bty="n", xlab = "", ylab
= "", col="red")
par(new=FALSE)
axis(side=4, at=pretty(range(AOSR$Production)))
mtext("Production (,000 tonnes)", side=4, line=3)
legend("bottomright", inset=.05, cex=1, bty="n", c("Area", "Production"),
lty=c(1, 1), lwd=c(1,1),col=c("blue", "red"))
```



```
par(mar=c(5, 4, 4, 6))
plot(AOSR$Year, AOSR$Area, type = "l", col="blue", bty="n", ylab="Area sown
(,000 ha)")
par(new=TRUE)
plot(AOSR$Year, AOSR$Yield, type="l", axes=FALSE, bty="n", xlab = "", ylab = "",
col="red")
par(new=FALSE)
axis(side=4, at=pretty(range(AOSR$Yield)))
mtext("Yield (t/ha)", side=4, line=3)
legend("bottomright", inset=.05, cex=1, bty="n", c("Area", "Yield"), lty=c(1,
1), lwd=c(1,1),col=c("blue", "red"))
```



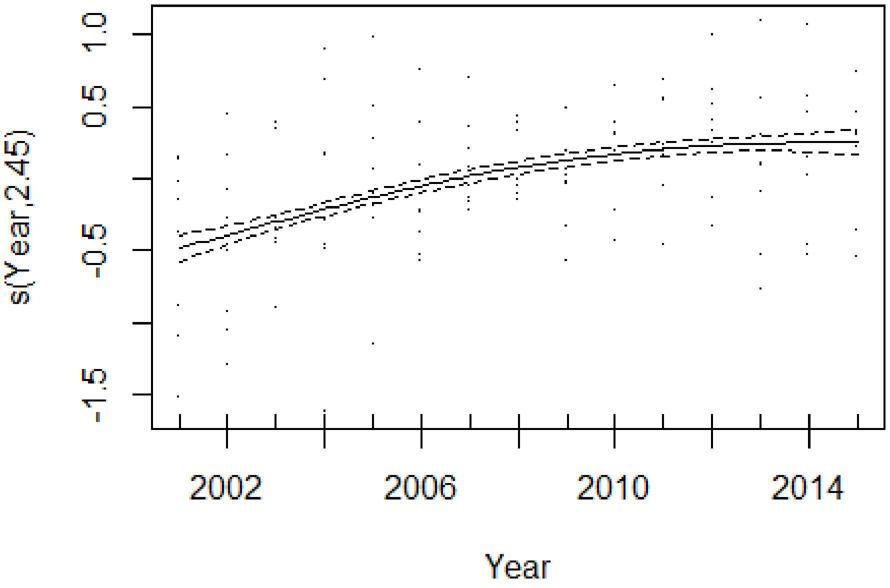
Poptrend analysis

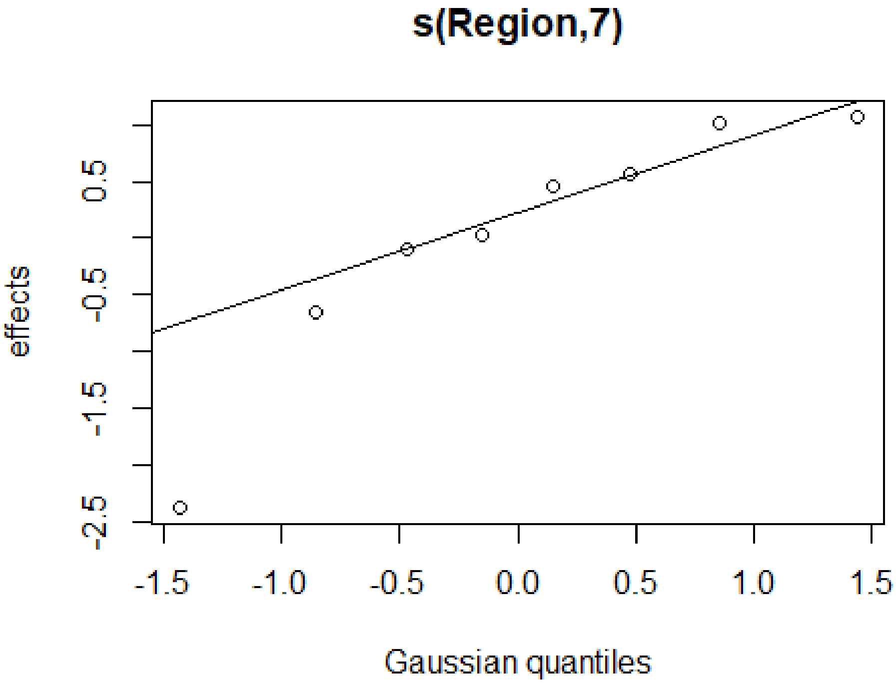
First, calculate theta for the dataset:

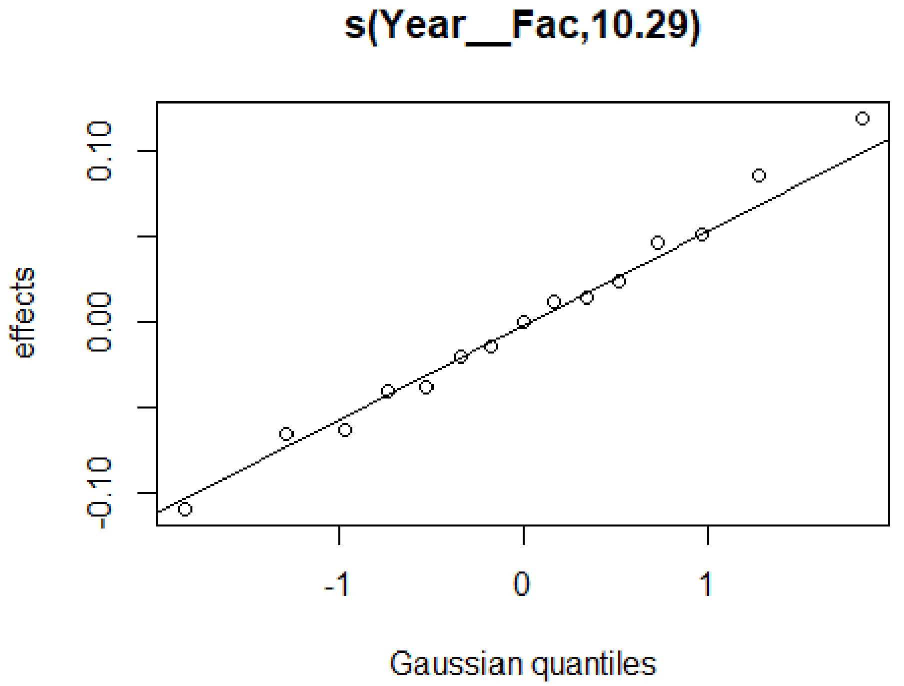
```
Trend3<-ptrend (Area ~ trend(Year, tempRE = TRUE, type = "smooth") + s(Region,
bs = "re"), family = nb(), data = OSR3)
Trend3$family$getTheta(TRUE) ## extract final theta estimate

## [1] 12460615
```

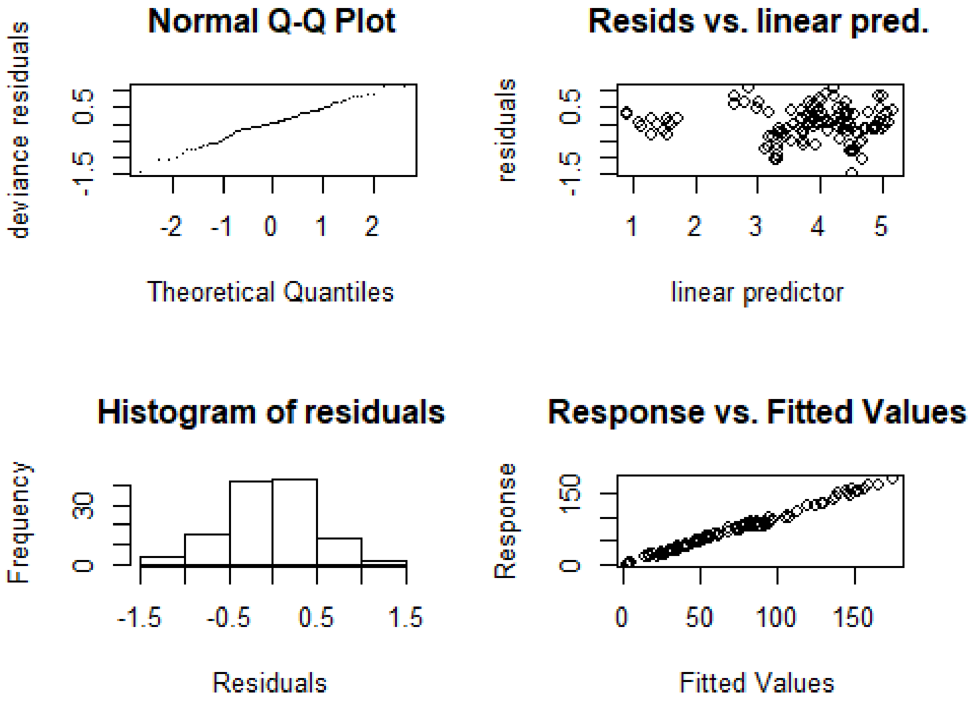
Then, run the poptrend analysis. Note that quasipoisson was selected as the best distribution for the oilseed rape cultivation data.

```
Trend<-ptrend (Area ~ trend(Year, tempRE = TRUE, type = "smooth") + s(Region, bs
= "re"), family = quasipoisson, data = OSR3)

checkFit(Trend)
```


s(Year,2.45)
1.0
0.5
-0.5
-1.5
2002
2006
2010
2014
Year


s(Region,7)
effects
0.5
-0.5
-1.5
-2.5
-1.5
-1.0
-0.5
0.0
0.5
1.0
1.5
Gaussian quantiles



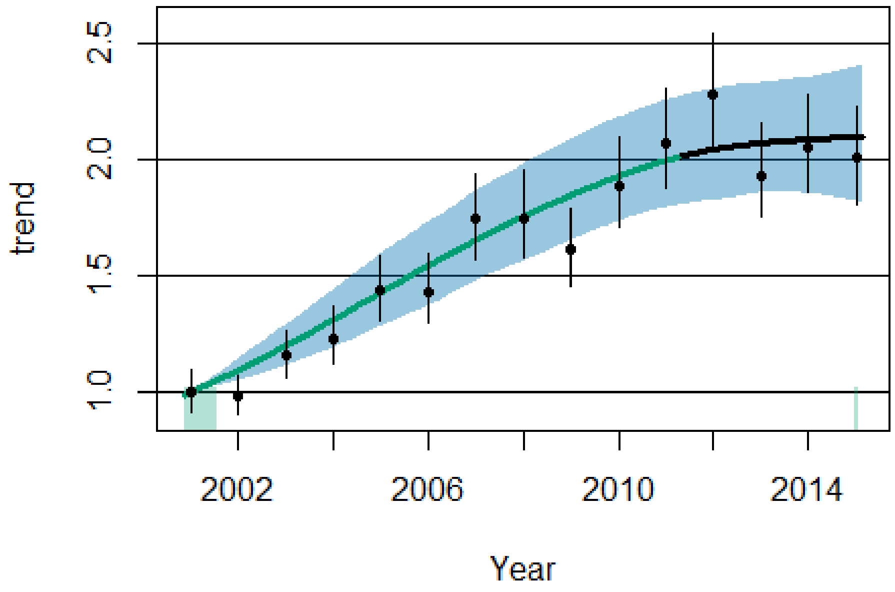


```
##
## Method: GCV   Optimizer: outer newton
## full convergence after 16 iterations.
## Gradient range [-3.433998e-09,3.220934e-07]
## (score 0.3271793 & scale 0.2688618).
## Hessian positive definite, eigenvalue range [1.014493e-05,0.006269579].
## Model rank =  33 / 33
```

```
##
## Basis dimension (k) checking results. Low p-value (k-index<1) may
## indicate that k is too low, especially if edf is close to k'.
##
##                   k'   edf k-index p-value
## s(Year)         9.00  2.45       1    0.46
## s(Region)       8.00  7.00      NA      NA
## s(Year__Fac) 15.00 10.29      NA      NA

#Check rate of change
change(Trend, 1987, 2017)

## Estimated percent change from Year = 2000.86 to 2015.14: 113% (84%, 145%)

#Summary
print(Trend)

##
## Family: quasipoisson
## Link function: log
##
## Formula: Area ~ trend(Year, tempRE = TRUE, type = "smooth") + s(Region,
##     bs = "re")
## Trend type:  smooth
```

Fig. 2.8

```
plot(Trend)
```



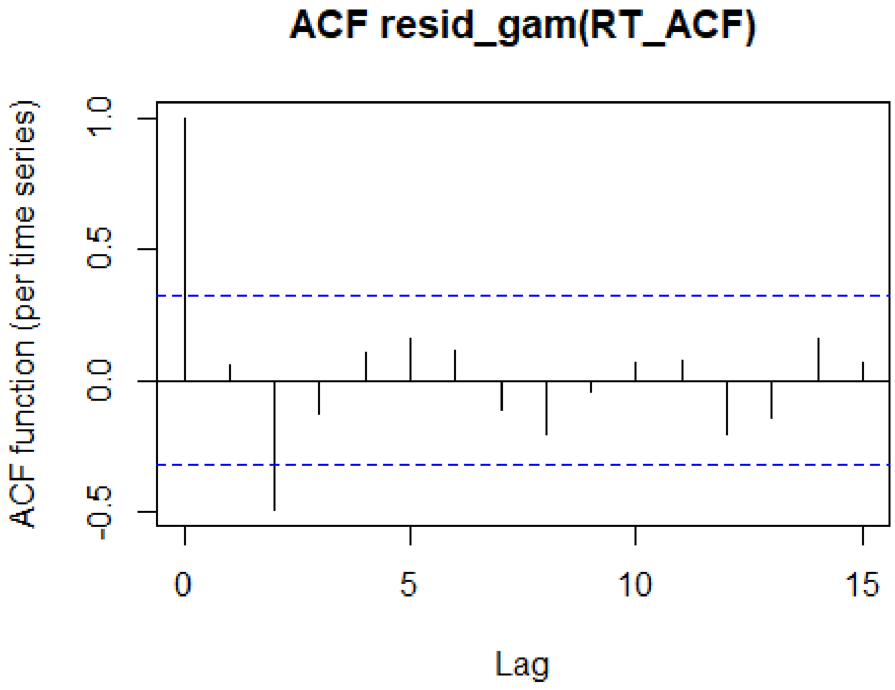
Long-term trends in pollen beetle abundance at Rothamsted and Broom's Barn

Section 2.4.3

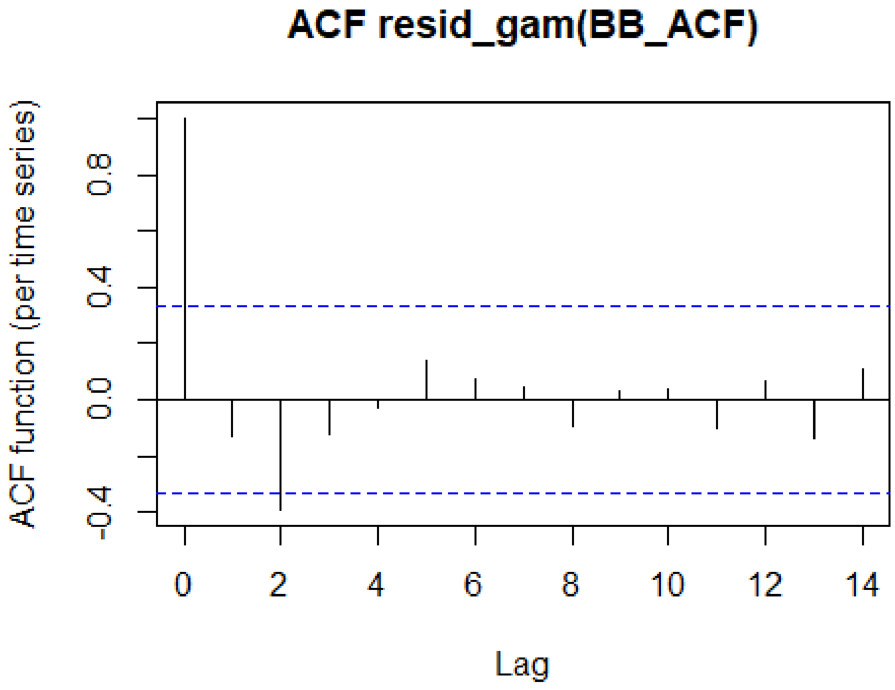
First, check for autocorrelation (Section 2.4.4):

Fig. 2.3

```
RT_ACF<-gam(RT ~ s(Year) +s(Area), family = nb(), method = "GCV.Cp", data = OSR,
rho=r1, AR.start= OSRUK$start.event)
acf_resid(RT_ACF)
```



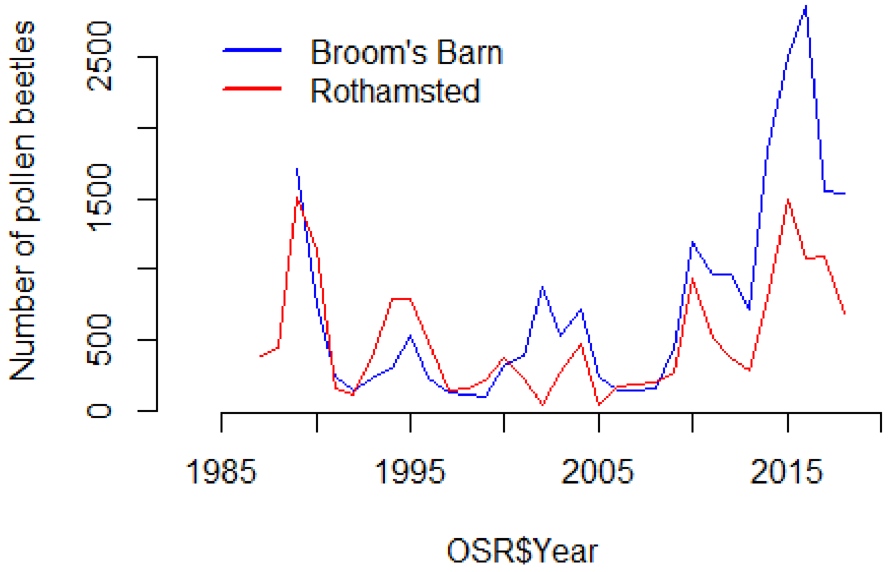
```
BB_ACF<-gam(BB ~ s(Year) +s(Area), family = nb(), method = "GCV.Cp", data = OSR,
rho=r1, AR.start= OSRUK$start.event)
acf_resid(BB_ACF)
```



Then plot the data.

Fig. 2.9

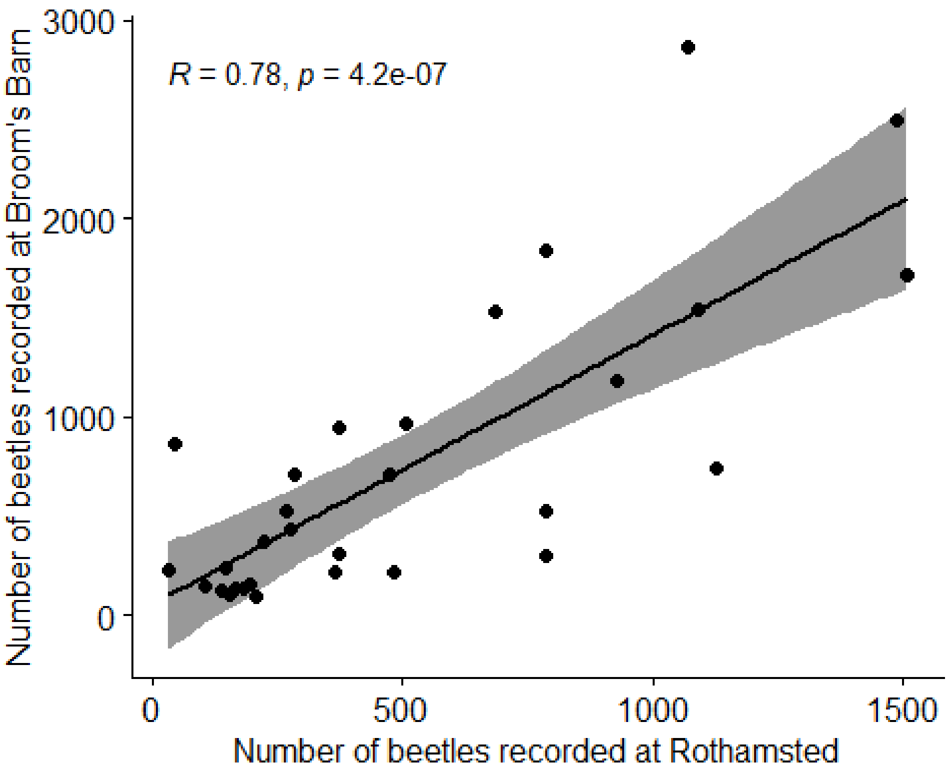
```
plot(OSR$Year, OSR$BB, type = "l", col="blue", bty="n", ylab="Number of pollen
beetles")
lines(OSR$Year, OSR$RT, type = "l", col="red", bty="n", ylab="Number of pollen
beetles")
legend("topleft", inset=.05, cex=1, bty="n", c("Broom's Barn", "Rothamsted"),
lty=c(1,1), lwd=c(2,2), col=c("blue", "red"))
```



Run a correlation test between Rothamsted and Broom’s Barn.

Fig. 2.10

```
ggscatter(OSR, x="RT", y = "BB", add = "reg.line", conf.int=TRUE, cor.coef=TRUE,
cor.method = "pearson", xlab = "Number of beetles recorded at Rothamsted", ylab
= "Number of beetles recorded at Broom's Barn")

## `geom_smooth()` using formula 'y ~ x'
```



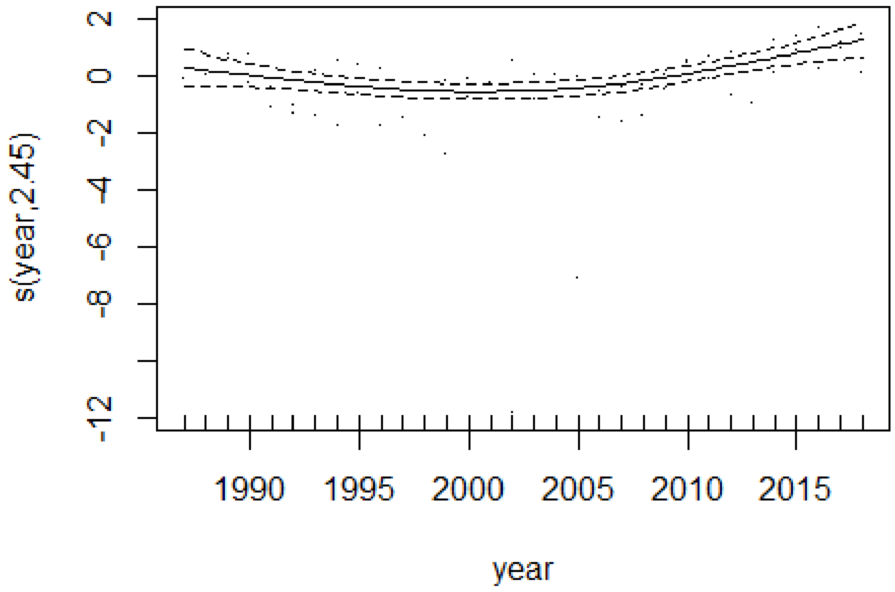
```
res1 <- cor.test(OSR$RT, OSR$BB,
                 method = "pearson")
res1

##
##  Pearson's product-moment correlation
##
## data:  OSR$RT and OSR$BB
## t = 6.549, df = 28, p-value = 4.233e-07
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.5801277 0.8889810
## sample estimates:
##       cor
## 0.777829
```

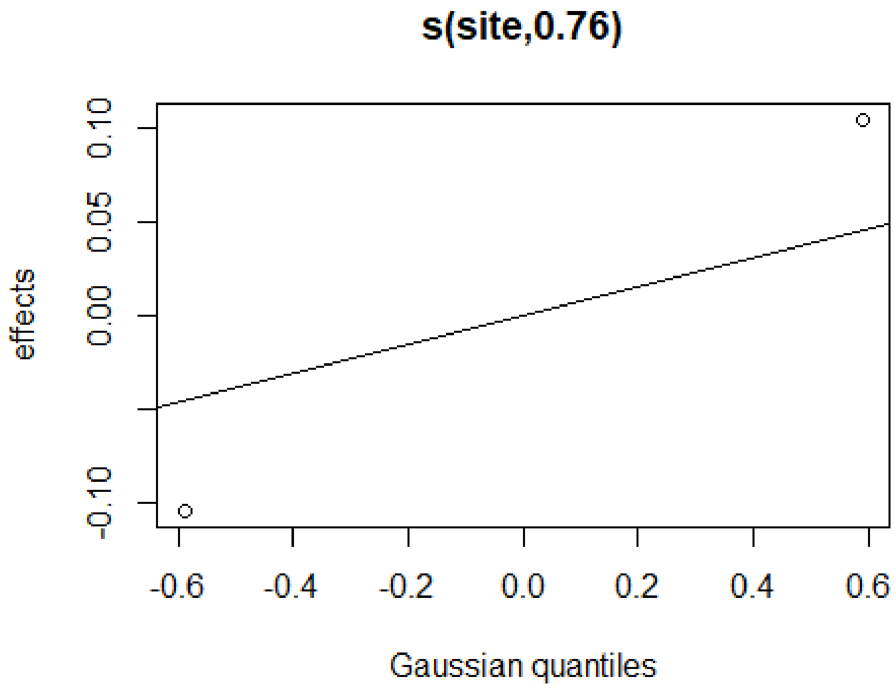
Model the trend

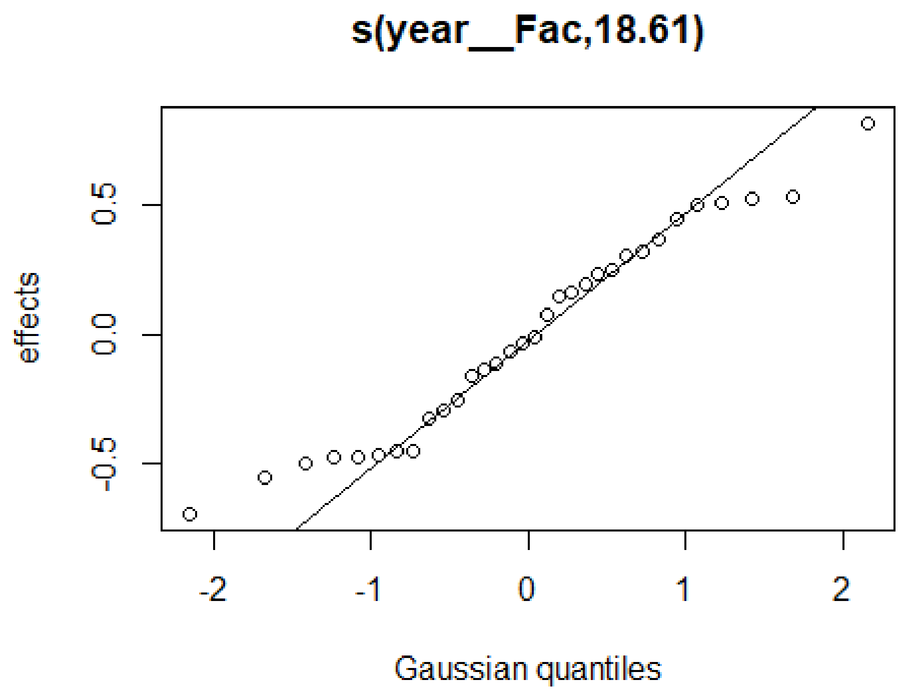
Fig. 2.11

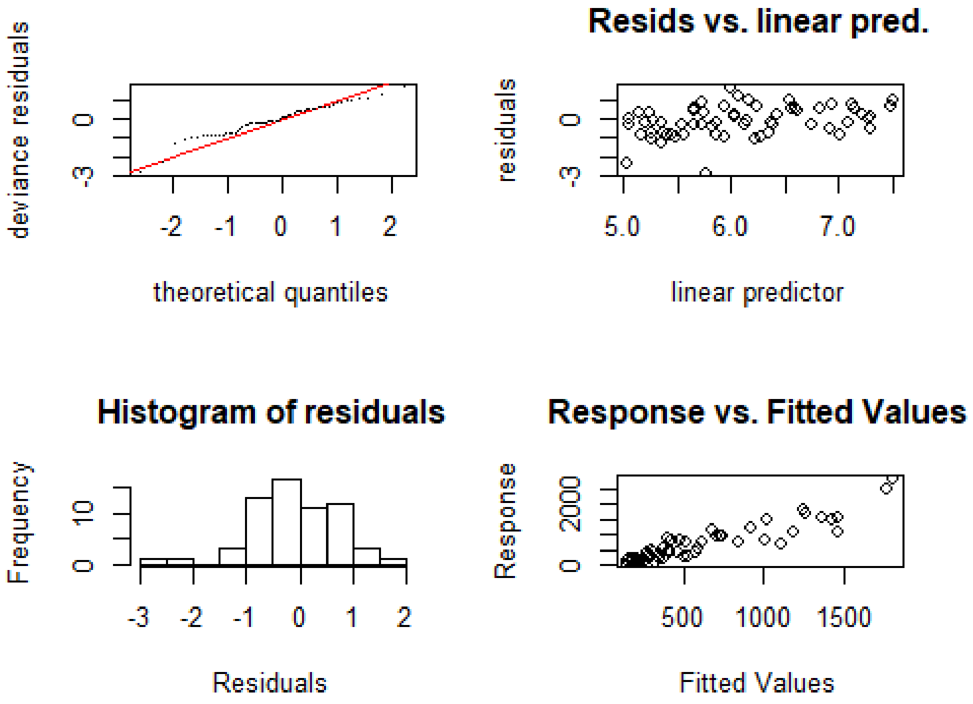
```
#Get theta
Trend2<-ptrend (count ~ trend(year, tempRE = TRUE, type = "smooth") + s(site, bs
= "re"), family = nb(), data = TA3)
Trend2$family$getTheta(TRUE) ## extract final theta estimate

## [1] 3.707052

#take result from above and add it between the brackets after negbin
Trends<-ptrend (count ~ trend(year, tempRE = TRUE, type = "smooth") + s(site, bs
= "re"), family = negbin(3.707052), data = TA3)
```

```
checkFit(Trends)
```



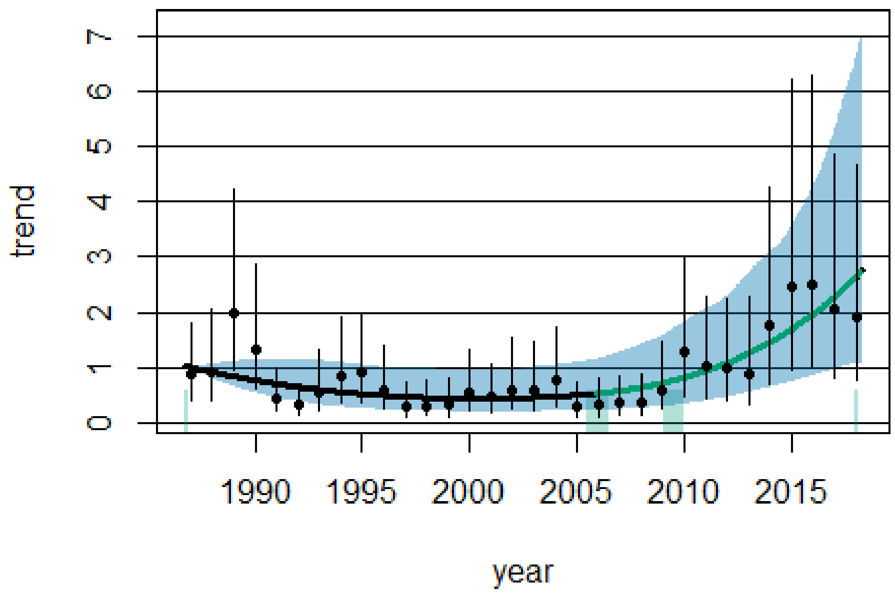






```
##
## Method: UBRE   Optimizer: outer newton
## full convergence after 4 iterations.
## Gradient range [3.32763e-09,1.080642e-06]
## (score 0.4065971 & scale 1).
## Hessian positive definite, eigenvalue range [0.004464057,0.137481].
## Model rank =  44 / 44
```

```
##
## Basis dimension (k) checking results. Low p-value (k-index<1) may
## indicate that k is too low, especially if edf is close to k'.
##
##                    k'    edf k-index p-value
## s(year)        9.000  2.452    1.41       1
## s(site)        2.000  0.759      NA      NA
## s(year__Fac) 32.000 18.613      NA      NA

plot(Trends)
```



```
change(Trends, 1987, 2018)

## Estimated percent change from year = 1987 to 2018: 162% (6.6%, 558%)

print(Trends)

##
## Family: Negative Binomial(3.707)
## Link function: log
##
## Formula: count ~ trend(year, tempRE = TRUE, type = "smooth") + s(site,
##     bs = "re")
## Trend type:  smooth
```

ptrend is a function that estimates a smooth or linear trend, count is the annual number of B. aeneus, trend is the function that sets up the trend component for ptrend, year is the year (1987 to 2018), tempRE=TRUE means that year is set as a random effect, type = “smooth” sets the trend to a smooth, rather than loglinear or index type, s(site, bs=”re”) sets the regression spline for site (here two sites, Rothamsted and Broom’s Barn) as a random effect, family =

negbin(3.707052) indicates the distribution is a negative binomial with Θ (the dispersion parameter) specified.

Seasonal trends

Re-run the analysis using seasonal data, this requires a separate dataset that divides the annual total catch into two generations (Spring (beetles caught before June 1st) and Summer (beetles caught after June 1st)).

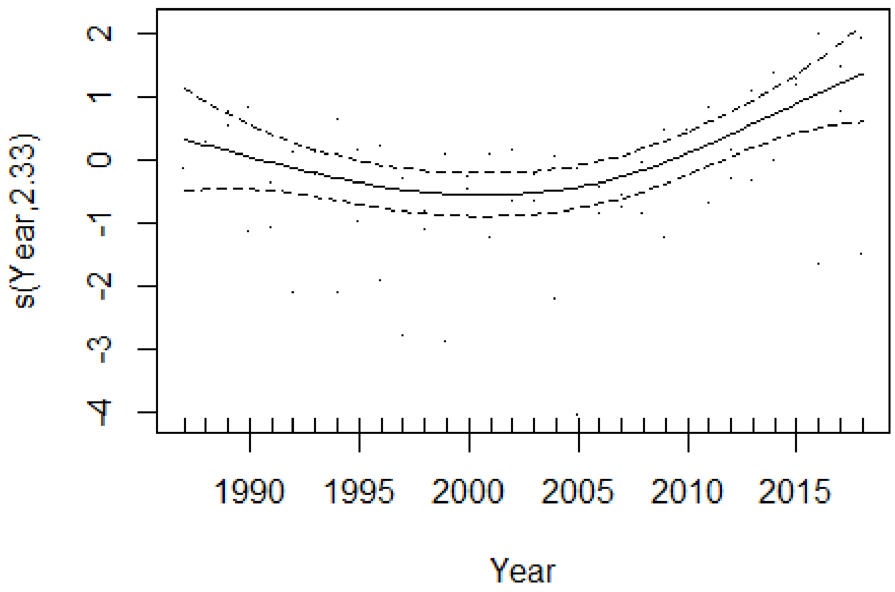
Monthly data are filtered from the main dataset.

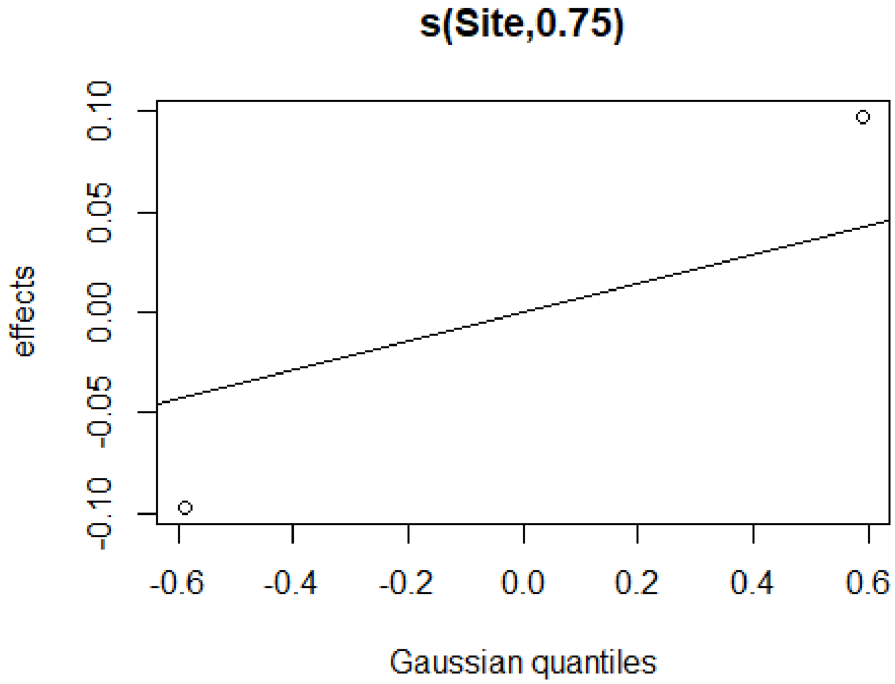
For the sake of brevity only the Spring analysis is presented here, the the other analyses are run with the same code, with theta calculated separately for each model and inserted manually into the code.

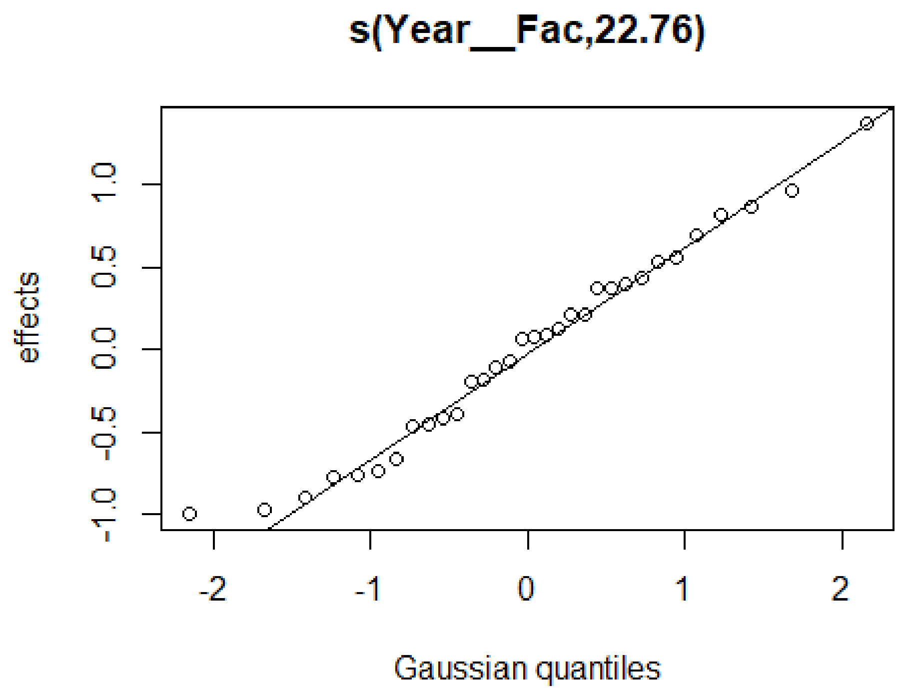
Fig. 2.12

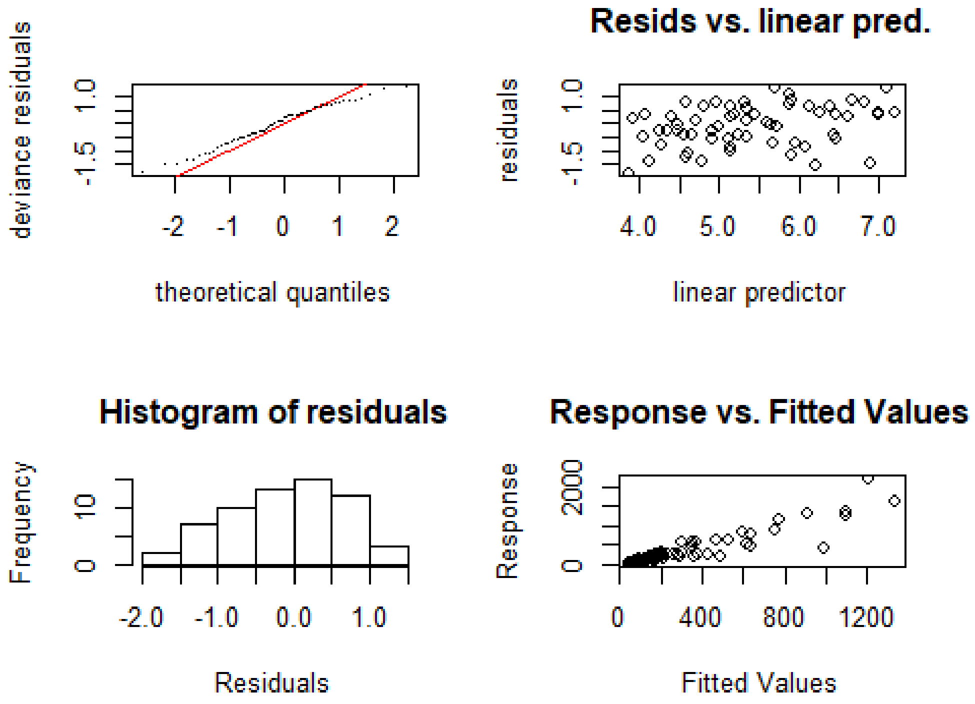
(Also, with alternative data for Figs. 2.13-2.18)

```
TrendSPRING<-ptrend (Count ~ trend(Year, tempRE = TRUE, type = "smooth") +
s(Site, bs = "re"), family = nb(), data = SPRING)
TrendSPRING$family$getTheta(TRUE) ## extract final theta estimate

## [1] 4.163855

#take result from above and add it between the brackets after negbin
TrendSPRING<-ptrend (Count ~ trend(Year, tempRE = TRUE, type = "smooth") +
s(Site, bs = "re"), family = negbin(3.961196), data = SPRING)
checkFit(TrendSPRING)
```


s(Year,2.33)
2
1
0
-1
-2
-3
-4
1990
1995
2000
2005
2010
2015
Year


s(Site,0.75)
effects
0.10
0.05
0.00
-0.05
-0.10
-0.6
-0.4
-0.2
0.0
0.2
0.4
0.6
Gaussian quantiles

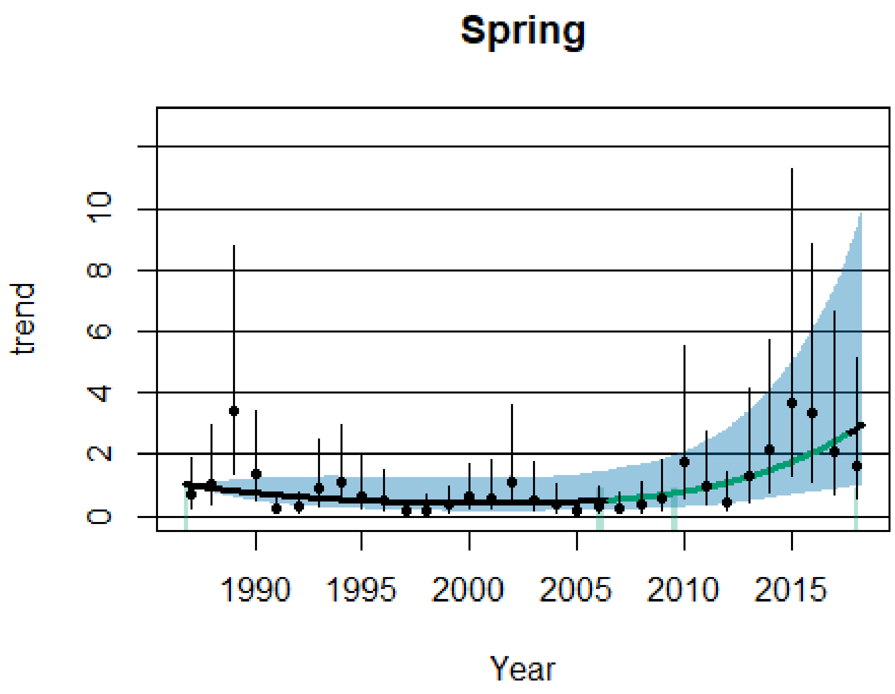




```
##
## Method: UBRE   Optimizer: outer newton
## full convergence after 5 iterations.
## Gradient range [5.444445e-10,4.694216e-08]
## (score 0.4421206 & scale 1).
## Hessian positive definite, eigenvalue range [0.004322224,0.122088].
## Model rank =  44 / 44
```

```
##
## Basis dimension (k) checking results. Low p-value (k-index<1) may
## indicate that k is too low, especially if edf is close to k'.
##
##                     k'    edf k-index p-value
## s(Year)          9.000  2.326    1.37       1
## s(Site)          2.000  0.751      NA      NA
## s(Year__Fac) 32.000 22.760      NA      NA
```

```
plot(TrendSPRING, main = "Spring")
```



```
change(TrendSPRING, 1987, 2018)
```

```
## Estimated percent change from Year = 1987 to 2018: 182% (-1.2%, 822%)
```

```
print(TrendSPRING)
```

```
##
## Family: Negative Binomial(3.961)
## Link function: log
##
## Formula: Count ~ trend(Year, tempRE = TRUE, type = "smooth") + s(Site,
##     bs = "re")
## Trend type:  smooth
```

Relationship between number of pollen beetles and area of OSR at different temporal and spatial scales.

Section 2.4.5

The relationship between number of pollen beetles and oilseed rape was invesigated using a GAM. First, calculate theta:

```
RTGAM<-gam(RT ~s(Area, bs="cs") + s(Year, bs="cs", k=15), family = nb(),
data=OSR, rho=r1, AR.start=OSR$start.event)
RTGAM$family$getTheta(TRUE)

## [1] 2.199612

BBGAM<-gam(BB ~s(Area, bs="cs") + s(Year, bs="cs", k=15), family = nb(),
data=OSR, rho=r1, AR.start=OSR$start.event)
BBGAM$family$getTheta(TRUE)

## [1] 12.70305
```

Then run the GAM using Area of OSR and Year as explanatory terms:

```
RT_resid<-gam(RT ~s(Area, bs="cs") + s(Year, bs="cs"), family =
negbin(theta=2.199612, link="log"), method = "REML", data=OSR)

BB_resid<-gam(BB ~s(Area, bs="cs") + s(Year, bs="cs"), family =
negbin(theta=12.70305, link="log"), method = "REML", data=OSR)

summary(RT_resid)

##
## Family: Negative Binomial(2.2)
## Link function: log
##
## Formula:
## RT ~ s(Area, bs = "cs") + s(Year, bs = "cs")
##
## Parametric coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)   6.0726     0.1195    50.8   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Approximate significance of smooth terms:
##            edf Ref.df Chi.sq  p-value
## s(Area) 0.8747      9  1.592    0.108
## s(Year) 2.5245      9 17.496 4.02e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

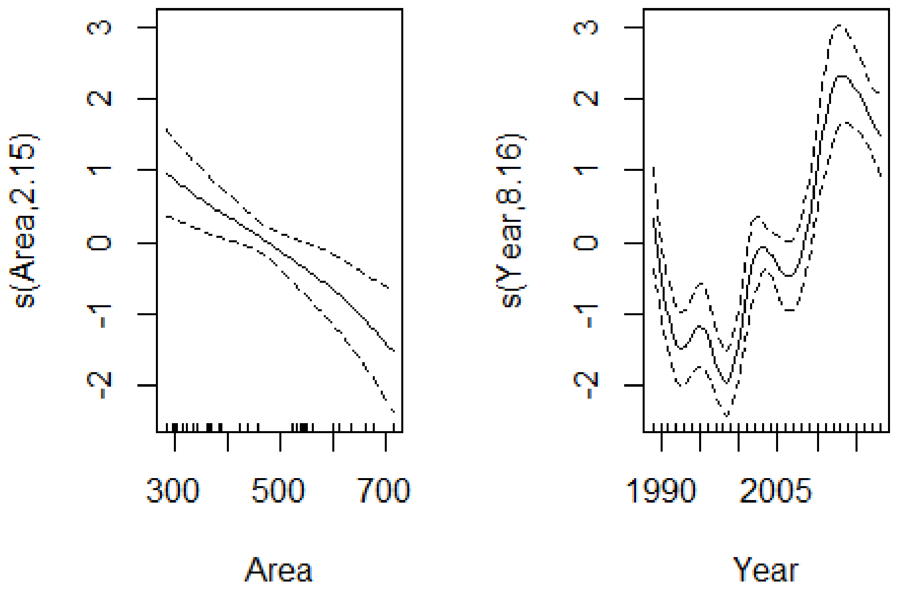
```
## R-sq.(adj) =  0.325   Deviance explained = 40.5%
## -REML = 227.77  Scale est. = 1         n = 32
```

```
summary(BB_resid)
```

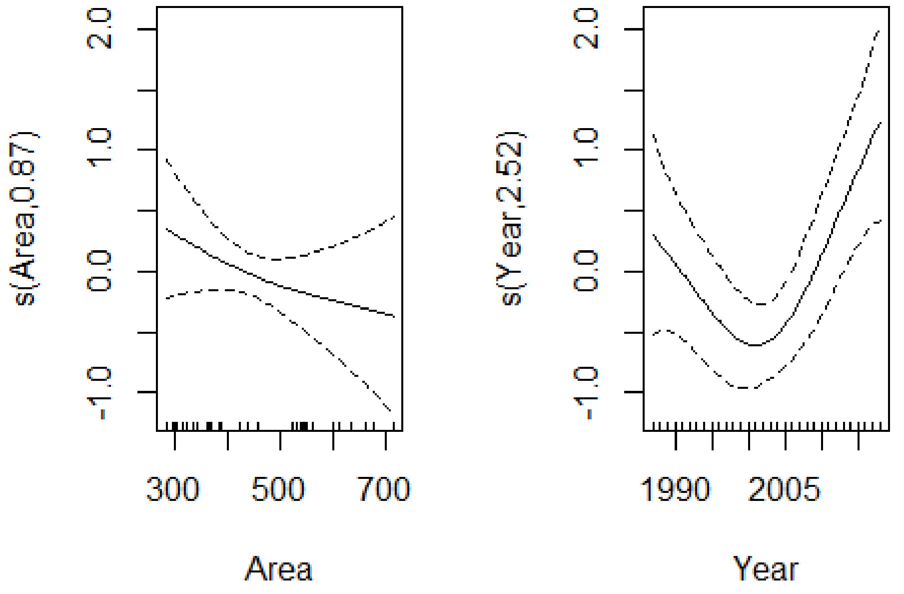
```
## 
## Family: Negative Binomial(12.703) 
## Link function: log 
## 
## Formula:
## BB ~ s(Area, bs = "cs") + s(Year, bs = "cs")
## 
## Parametric coefficients:
##             Estimate Std. Error z value Pr(>|z|)    
## (Intercept)   6.1880     0.0522   118.5   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
## Approximate significance of smooth terms:
##           edf Ref.df Chi.sq  p-value    
## s(Area) 2.147      9  15.25 4.69e-05 ***
## s(Year) 8.157      9 253.34  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
## R-sq.(adj) =  0.816   Deviance explained = 92.3%
## -REML = 218.55  Scale est. = 1         n = 30
```

Figs 2.19 & 2.20

```
par(mfrow= c(1,2))
plot(BB_resid)
```



```
plot(RT_resid)
```

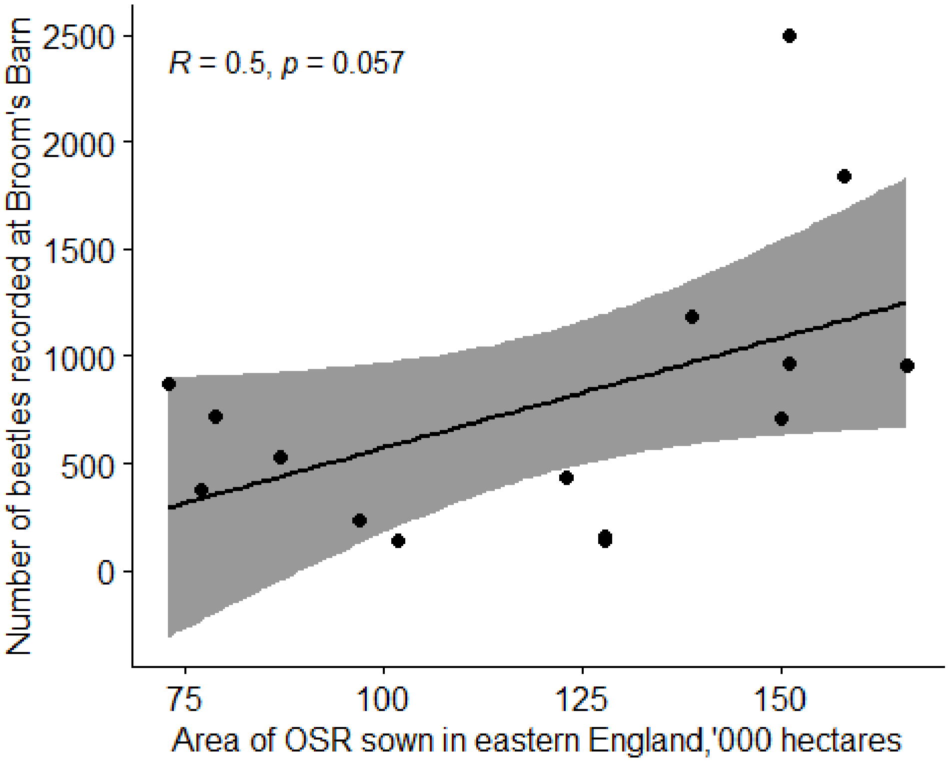


```
par(mfrow=c(1,1))
```

For shorter-term, larger spatial-scale comparisons the data were not sufficient for a GAM, thus Pearson correlations were calculated.

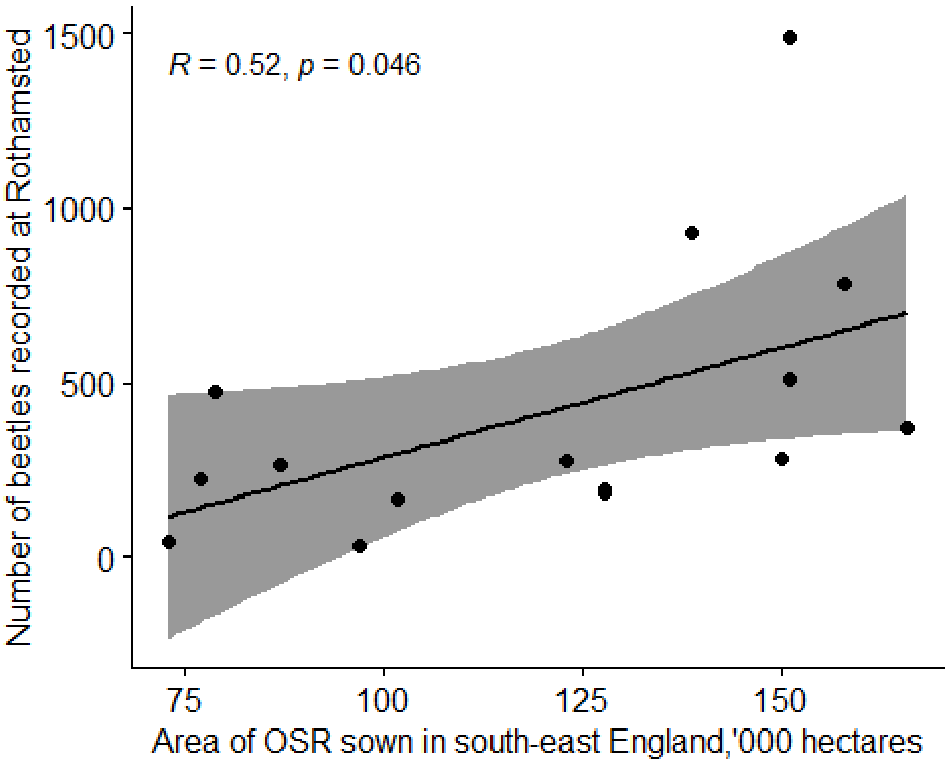
Figs 2.21 & 2.22

```
ggscatter(OSR4, x="Area", y = "BB", add = "reg.line", conf.int=TRUE,
cor.coef=TRUE, cor.method = "pearson", xlab = "Area of OSR sown in eastern
England,'000 hectares", ylab = "Number of beetles recorded at Broom's Barn")

## `geom_smooth()` using formula 'y ~ x'
```



```
resBB <- cor.test(OSR4$Area, OSR4$BB,
                  method = "pearson")
resBB

##
##  Pearson's product-moment correlation
##
## data:  OSR4$Area and OSR4$BB
## t = 2.0865, df = 13, p-value = 0.0572
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  -0.01533395  0.80626079
## sample estimates:
##       cor
## 0.5008632
```

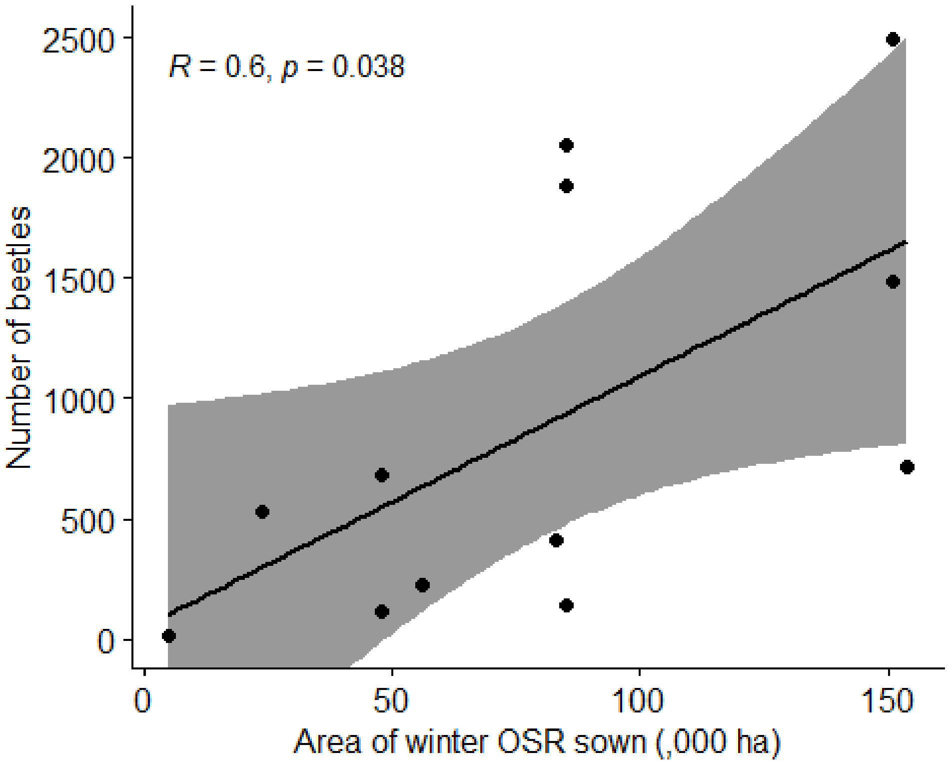
```
ggscatter(OSR4, x="Area", y = "RT", add = "reg.line", conf.int=TRUE,
cor.coef=TRUE, cor.method = "pearson", xlab = "Area of OSR sown in south-east
England,'000 hectares", ylab = "Number of beetles recorded at Rothamsted")

## `geom_smooth()` using formula 'y ~ x'
```



```
resRT <- cor.test(OSR4$Area, OSR4$RT,
                  method = "pearson")
resRT

##
##  Pearson's product-moment correlation
##
## data:  OSR4$Area and OSR4$RT
## t = 2.205, df = 13, p-value = 0.04607
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.01290969 0.81592254
## sample estimates:
##       cor
## 0.5217223
```

Fig. 2.23

```
ggscatter(OSR5, x="Area.Sown", y = "X2015.total", add = "reg.line",
conf.int=TRUE, cor.coef=TRUE, cor.method = "pearson", xlab = "Area of winter OSR
sown (,000 ha)", ylab = "Number of beetles", ylim=c(0, 2500))

## `geom_smooth()` using formula 'y ~ x'
```



```
res5 <- cor.test(OSR5$Area.Sown, OSR5$X2015.total,
                method = "pearson")
res5

##
##  Pearson's product-moment correlation
##
## data:  OSR5$Area.Sown and OSR5$X2015.total
## t = 2.3924, df = 10, p-value = 0.0378
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.04502568 0.87445366
## sample estimates:
##       cor
## 0.6033369
```

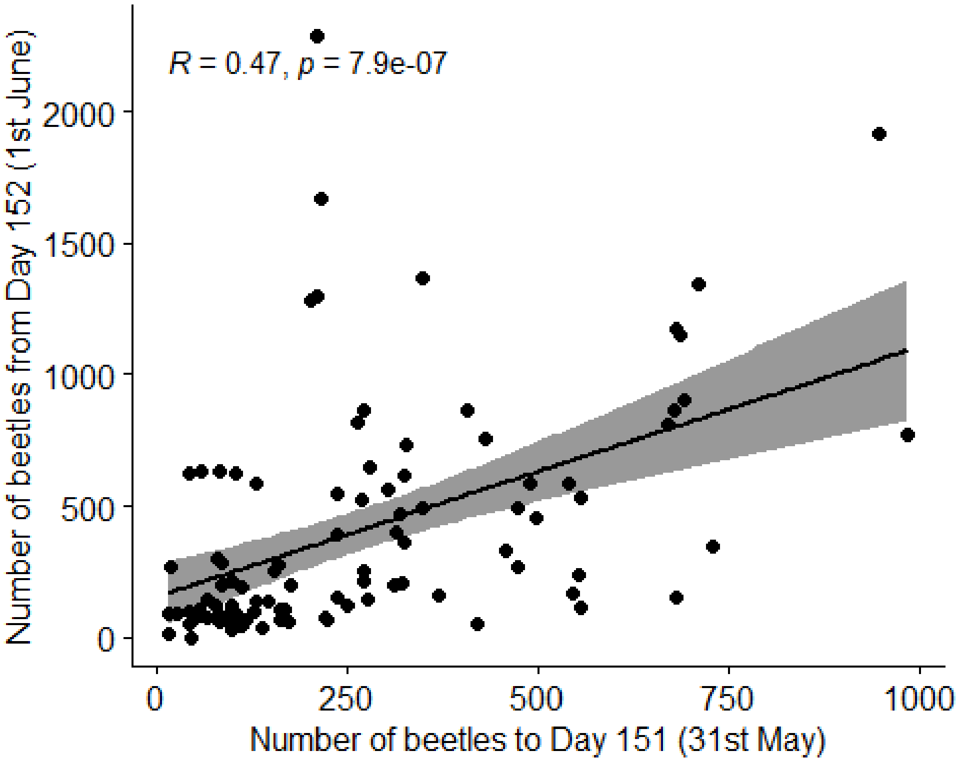
Relationship between number of pollen beetles caught in spring and summer

Section 2.4.6

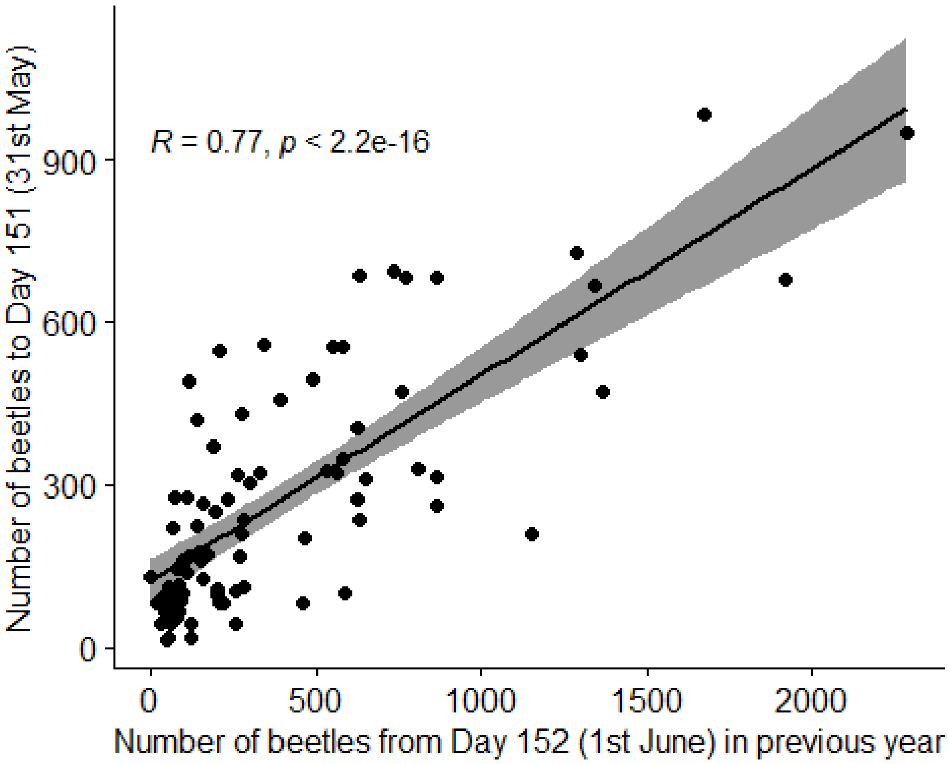
Here, the two generations of pollen beetle are investigated to see if counts of one can predict counts of the following. In addition, data for beetles entering hibernation are used to forecast the abundance of those exiting hibernation the following year.

Figs. 2.24 & 2.25

```
Trap<-factor(PB$Site)

ggscatter(PB, x="Spring", y = "Summer", add = "reg.line", conf.int=TRUE,
cor.coef=TRUE, cor.method = "pearson", xlab = "Number of beetles to Day 151
(31st May)", ylab = "Number of beetles from Day 152 (1st June)")

## `geom_smooth()` using formula 'y ~ x'
```



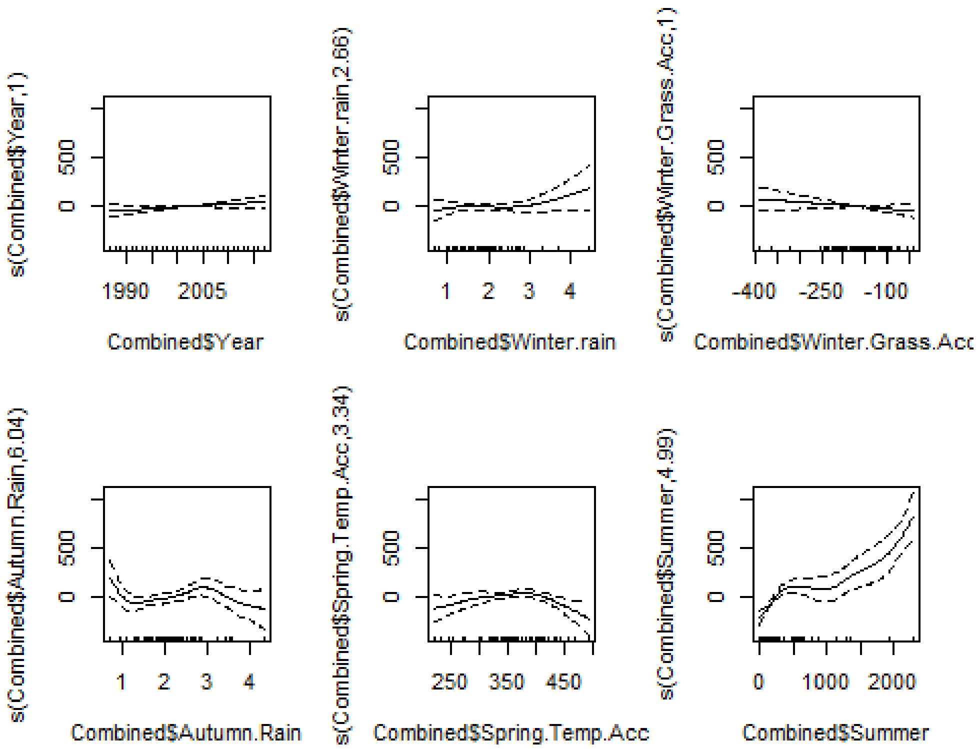
```
res <- cor.test(PB$Spring, PB$Summer,
                 method = "pearson")
res

##
##  Pearson's product-moment correlation
##
## data:  PB$Spring and PB$Summer
## t = 5.2832, df = 96, p-value = 7.92e-07
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.3049088 0.6151090
## sample estimates:
##       cor
## 0.4746152

ggscatter(PB, x="Summer", y = "Overwinter", add = "reg.line", conf.int=TRUE,
cor.coef=TRUE, cor.method = "pearson", xlab = "Number of beetles from Day 152
(1st June) in previous year", ylab = "Number of beetles to Day 151 (31st May)")

## `geom_smooth()` using formula 'y ~ x'
```



```
res2 <- cor.test(PB$Summer, PB$Overwinter,
                 method = "pearson")
res2

##
##  Pearson's product-moment correlation
##
## data:  PB$Summer and PB$Overwinter
## t = 11.108, df = 85, p-value < 2.2e-16
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.6669439 0.8434235
## sample estimates:
##       cor
## 0.7694866
```

After model selection the model defined as model1, was selected, gam is the GAM function in mgcv, Overwinter is the overwintered generation active in spring with all the explanatory parameters defined with smooth terms s(x). Due to the lack of replication – only two sites were used – it was not possible to include any random effects of, for example, year or site.

```
par(mfrow=c(2,3))
model1<-
gam(Combined$Overwinter~s(Combined$Year)+s(Combined$Winter.rain)+s(Combined$Wint
er.Grass.Acc)+s(Combined$Autumn.Rain)+s(Combined$Spring.Temp.Acc)+s(Combined$Sum
mer))
plot(model1)
```

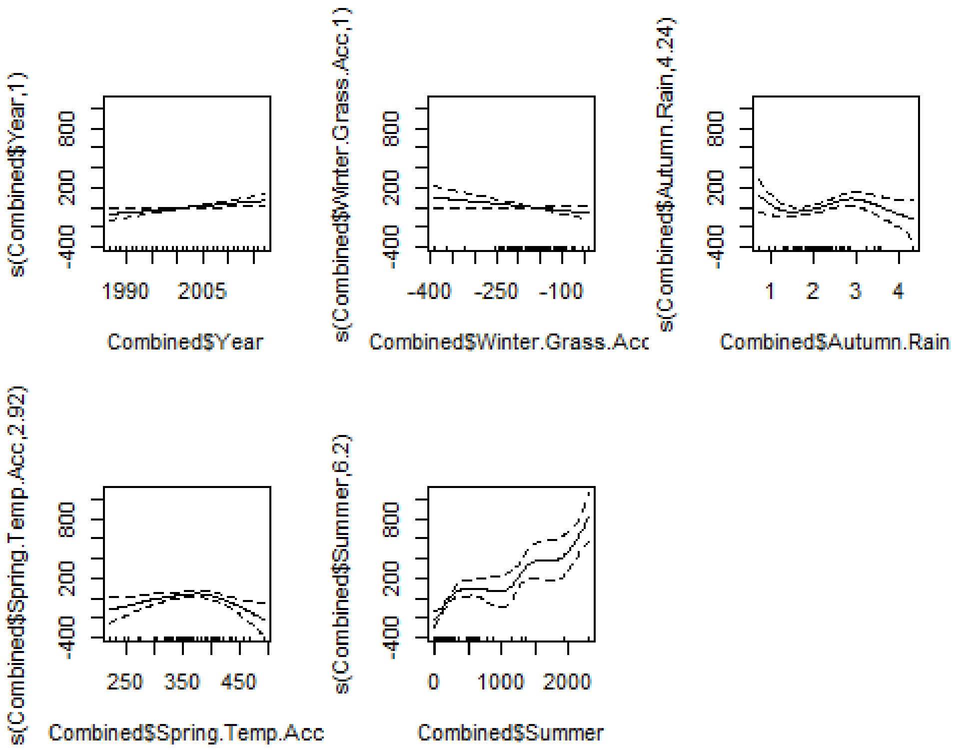


```
summary(model1)

##
## Family: gaussian
## Link function: identity
##
## Formula:
## Combined$Overwinter ~ s(Combined$Year) + s(Combined$Winter.rain) +
##     s(Combined$Winter.Grass.Acc) + s(Combined$Autumn.Rain) +
##     s(Combined$Spring.Temp.Acc) + s(Combined$Summer)
##
## Parametric coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   241.35      13.56    17.8   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Approximate significance of smooth terms:
##                                edf Ref.df      F p-value
## s(Combined$Year)             1.000  1.000  2.105  0.1546
## s(Combined$Winter.rain)      2.657  3.307  0.917  0.4634
## s(Combined$Winter.Grass.Acc) 1.000  1.000  1.606  0.2124
## s(Combined$Autumn.Rain)      6.041  7.043  1.991  0.0775 .
## s(Combined$Spring.Temp.Acc)  3.343  4.109  2.899  0.0340 *
## s(Combined$Summer)           4.988  5.889 17.294  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## R-sq.(adj) =  0.748   Deviance explained = 82.9%
## GCV =  16556  Scale est. = 11029     n = 60
```

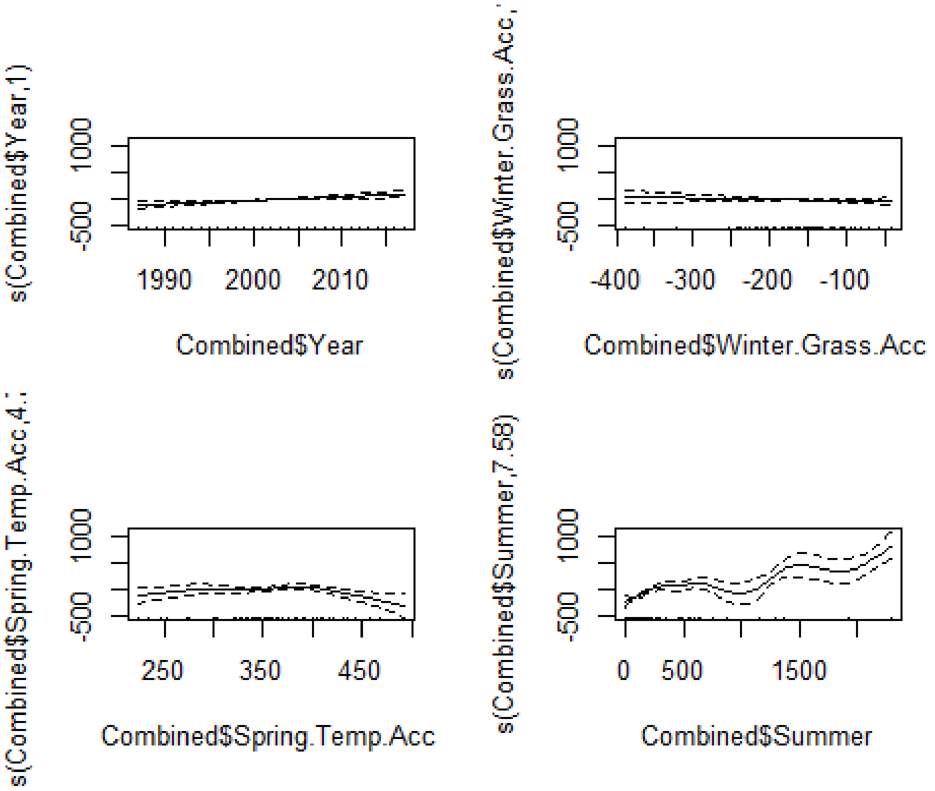
```
#Winter rainfall least significant
model2<-
gam(Combined$Overwinter~s(Combined$Year)+s(Combined$Winter.Grass.Acc)+s(Combined
$Autumn.Rain)+s(Combined$Spring.Temp.Acc)+s(Combined$Summer))
plot(model2)
summary(model2)
```

```
##
## Family: gaussian
## Link function: identity
##
## Formula:
## Combined$Overwinter ~ s(Combined$Year) + s(Combined$Winter.Grass.Acc) +
##     s(Combined$Autumn.Rain) + s(Combined$Spring.Temp.Acc) +
s(Combined$Summer)
##
## Parametric coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   241.35      13.96   17.29   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Approximate significance of smooth terms:
##                                edf Ref.df      F p-value
## s(Combined$Year)             1.000  1.000  4.573  0.0381 *
## s(Combined$Winter.Grass.Acc) 1.000  1.000  2.889  0.0962 .
## s(Combined$Autumn.Rain)      4.243  5.193  1.942  0.1075
## s(Combined$Spring.Temp.Acc)  2.920  3.627  3.325  0.0222 *
## s(Combined$Summer)           6.197  7.168 15.387  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## R-sq.(adj) =  0.733   Deviance explained = 80.3%
## GCV =  16073  Scale est. = 11690     n = 60
```

```
par(mfrow=c(2,2))
```



```
model3<-
gam(Combined$Overwinter~s(Combined$Year)+s(Combined$Winter.Grass.Acc)+s(Combined
$Spring.Temp.Acc)+s(Combined$Summer))
plot(model3)
```



```
summary(model3)
```

```
##
## Family: gaussian
## Link function: identity
##
## Formula:
## Combined$Overwinter ~ s(Combined$Year) + s(Combined$Winter.Grass.Acc) +
##     s(Combined$Spring.Temp.Acc) + s(Combined$Summer)
##
## Parametric coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   241.35      14.47   16.68   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Approximate significance of smooth terms:
##                                edf Ref.df      F p-value
## s(Combined$Year)             1.000  1.000 10.307 0.00245 **
## s(Combined$Winter.Grass.Acc) 1.000  1.000  0.769 0.38527
## s(Combined$Spring.Temp.Acc)  4.752  5.777  2.498 0.03824 *
## s(Combined$Summer)           7.575  8.418 13.874 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## R-sq.(adj) =  0.713   Deviance explained = 78.3%
## GCV =  16872  Scale est. = 12562     n = 60
```

```
model4<-
gam(Combined$Overwinter~s(Combined$Year)+s(Combined$Spring.Temp.Acc)+s(Combined$
Summer))
plot(model4)
summary(model4)
```

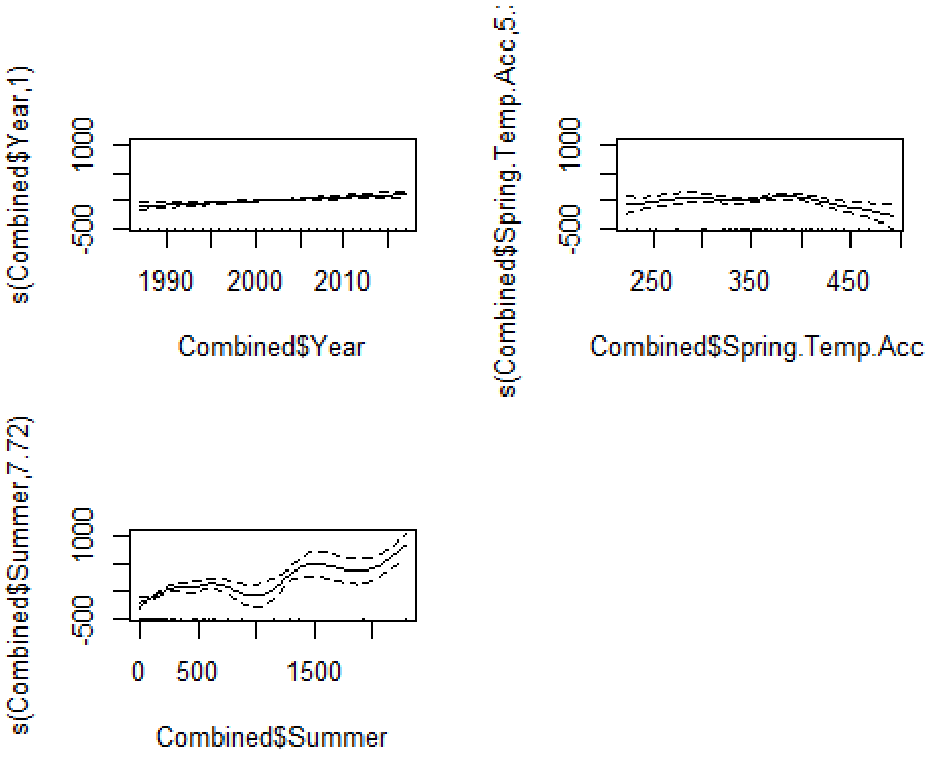
```
##
## Family: gaussian
## Link function: identity
##
## Formula:
## Combined$Overwinter ~ s(Combined$Year) + s(Combined$Spring.Temp.Acc) +
##     s(Combined$Summer)
##
## Parametric coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   241.35      14.27   16.91   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Approximate significance of smooth terms:
##                               edf Ref.df      F p-value
## s(Combined$Year)            1.000  1.000 11.455 0.00149 **
## s(Combined$Spring.Temp.Acc) 5.308  6.382  2.402 0.04089 *
## s(Combined$Summer)          7.716  8.513 14.301 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

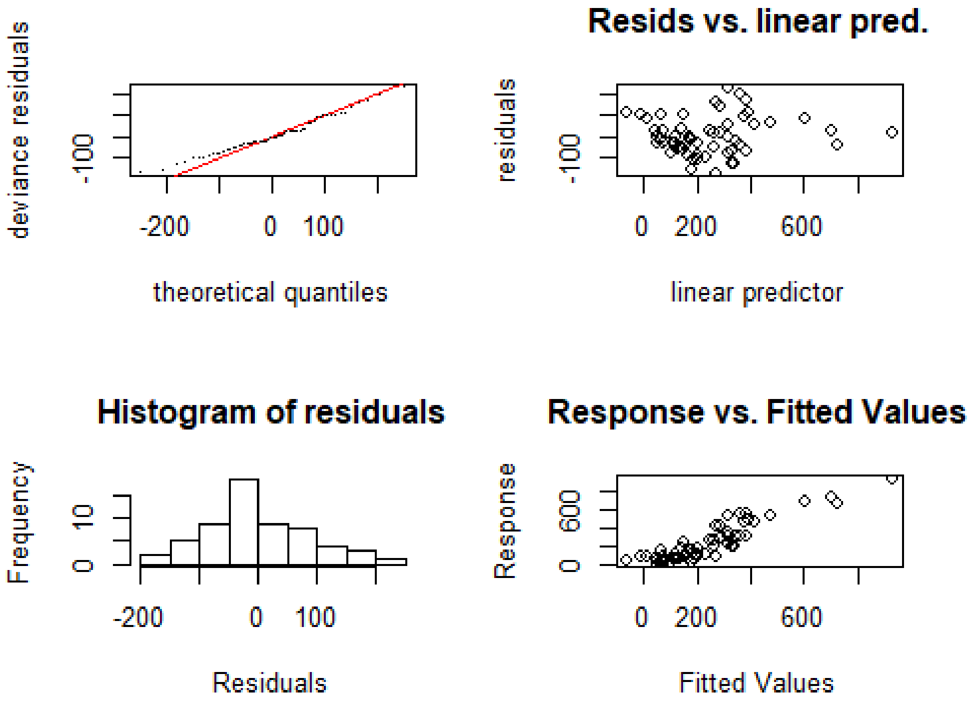
```
## R-sq.(adj) =  0.721   Deviance explained = 78.7%
## GCV =  16308  Scale est. = 12225     n = 60
```

```
AIC(model1, model2, model3, model4)
```

```
##              df      AIC
## model1 21.03009 746.4562
## model2 17.35993 747.8812
## model3 16.32746 751.5339
## model4 16.02381 749.7007
```

```
gam.check(model1)
```

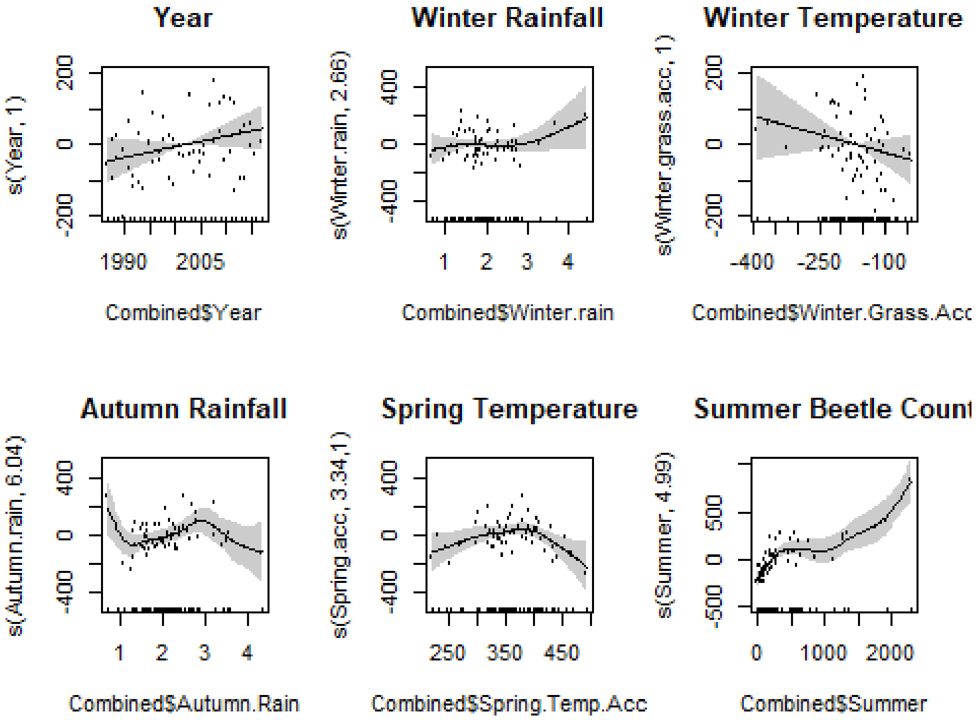




```
##
## Method: GCV   Optimizer: magic
## Smoothing parameter selection converged after 24 iterations.
## The RMS GCV score gradient at convergence was 0.000337139 .
## The Hessian was positive definite.
## Model rank =  55 / 55
##
```

```
## Basis dimension (k) checking results. Low p-value (k-index<1) may
## indicate that k is too low, especially if edf is close to k'.
##
##                                k'  edf k-index p-value
## s(Combined$Year)             9.00 1.00    1.20    0.95
## s(Combined$Winter.rain)      9.00 2.66    1.02    0.49
## s(Combined$Winter.Grass.Acc) 9.00 1.00    1.01    0.50
## s(Combined$Autumn.Rain)      9.00 6.04    1.17    0.92
## s(Combined$Spring.Temp.Acc)  9.00 3.34    1.17    0.85
## s(Combined$Summer)           9.00 4.99    1.22    0.93
```

```
par(mfrow=c(2,3))

plot.gam(model1,select =1, residuals=T,se=T,pch=1, cex=0.1, all.terms=T,
ylim=c(-200,200), shade=TRUE,  main="Year", ylab = "s(Year, 1)")

plot.gam(model1, select=2, residuals=T, se=T, pch=1, cex=0.1, all.terms=T,
ylim=c(-500, 500), shade=TRUE, main="Winter Rainfall", ylab="s(Winter.rain,
2.66)")


plot.gam(model1,select =3, residuals=T,se=T,pch=1, cex=0.1, all.terms=T,
ylim=c(-200,200), shade=TRUE,  main="Winter Temperature", ylab =
"s(Winter.grass.acc, 1)")

plot.gam(model1, select=4, residuals=T, se=T, pch=1, cex=0.1, all.terms=T,
ylim=c(-500, 500), shade=TRUE, main="Autumn Rainfall", ylab="s(Autumn.rain,
6.04)")

plot.gam(model1,select =5, residuals=T,se=T,pch=1, cex=0.1, all.terms=T,
ylim=c(-500,500), shade=TRUE,  main="Spring Temperature", ylab = "s(Spring.acc,
3.34,1)")

plot.gam(model1, select=6, residuals=T, se=T, pch=1, cex=0.1, all.terms=T,
ylim=c(-500, 1000), shade=TRUE, main="Summer Beetle Count", ylab="s(Summer,
4.99)")
```



```
par(mfrow=c(1,1))
```

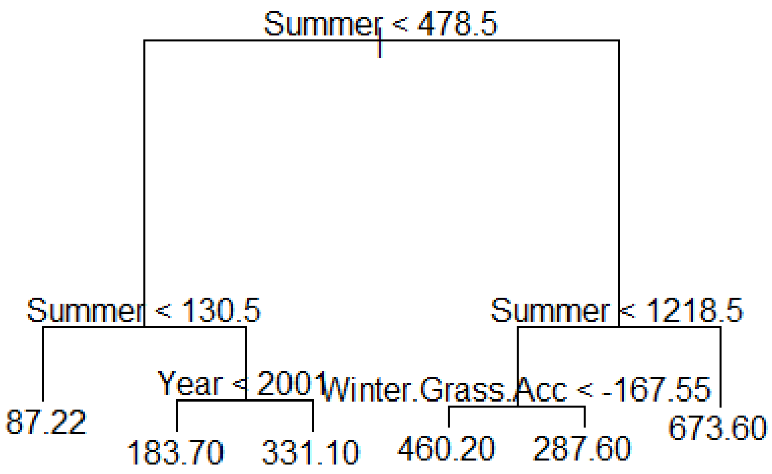
Drivers of overwintered beetle abundance

Section 2.4.7

The tree package was used to explore the autumn and winter environmental drivers of the abundance of beetles emerging from hibernation.

Fig. 2.27

```
modelt<-tree(Overwinter~.,data=Combined)
plot(modelt)
text(modelt)
```


Summer < 478.5
Summer < 130.5
Summer < 1218.5
Year < 2001
Winter.Grass.Acc < -167.55
87.22
183.70
331.10
460.20
287.60
673.60

Appendix A, Chapter 3

This chapter used a range of analytical methods to interrogate the daily data from the suction-traps. These data proved challenging, being zero-inflated and exhibiting day-to-day changes of pollen beetle abundance of two or more orders of magnitude. This meant that no statistical approach was wholly satisfactory, but some methods were more able to account for these challenges than others.

Libraries

The following R packages are required for the analyses in this chapter:

```r
library("mgcv")

library("fitdistrplus")

library("akima")

library("lme4")

library("mcmc")

library("deSolve")

library("spdep")

library("spatstat")

library("tree")

library("abind")
library("keras")

library("reticulate")

library("GSIF")

library("raster")

library("SearchTrees")
library("MASS")
library("psych")

library("pscl")

library("corrr")

library("AER")

library("broom")

library("performance")

library("sp")
library("ROCR")
```

```
library("randomForest")

library("poptrend")

library("gratia")

library("plotrix")

library("ggeffects")

library("tensorflow")

library("chron")

library("tidyverse")

library("magrittr")

library("dplyr")
library("tidyr")
library("tibble")

library("ggplot2")
library("RColorBrewer")
library("cartography")

library("gratia")
library("ggfortify")

library("scales")

library("factoextra")

library("rpart")
library("partykit")
```

Packages chron, tidyverse, magrittr, dplyr, tidyr and tibble were used to rearrange or prepare data where required for analyses, ggplot2, RColorBrewer, cartography, gratia, ggfortify, scales, factoextra, rpart and rpartykit were used for graphical output, the rest were required for the analyses. Note that the keras and reticulate packages require the Python machine learning software "tensorflow" to be installed (see: https://www.tensorflow.org/) plus the tensorflow R package.

Data

The following files were used, all are available on request, as are the original R code files:

```
setwd("P:/Final datasets")
daily<-read.csv("master dataset.csv")
data <- read.csv("insectAndMeteo_12traps_1987to2018_new_z_sum.csv")
positivedaily<-read.csv("master dataset nonnegative.csv")
```

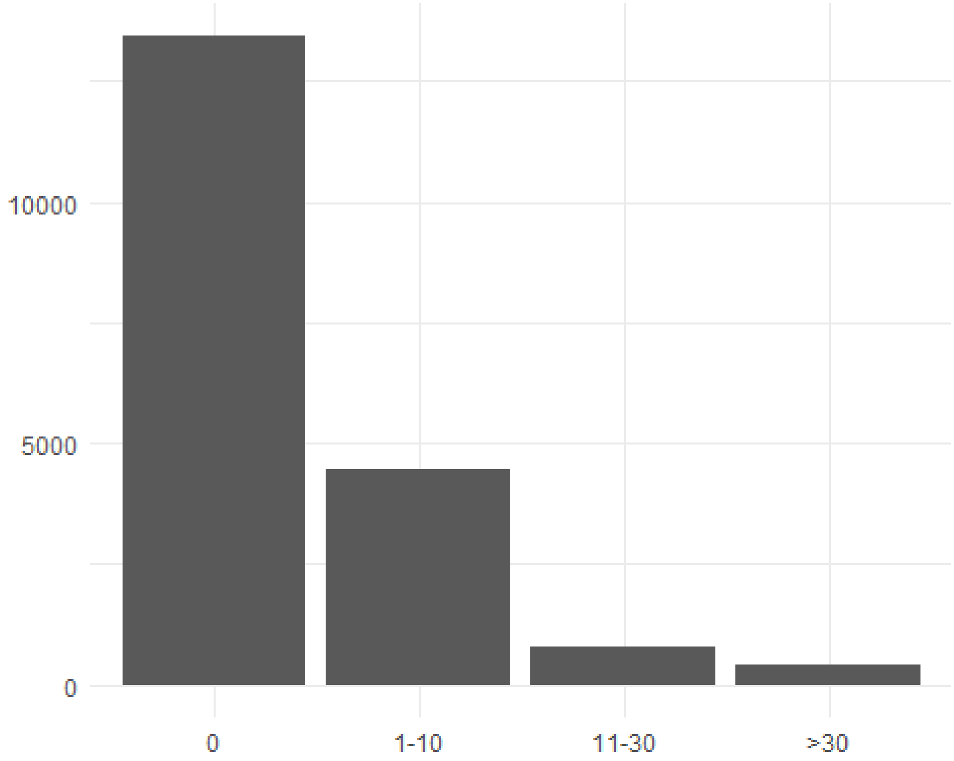
Statistical distribution of the data

All datasets were interrogated using the fitdistrplus package to determine their statistical distribution. See Appendix A, Chapter 2 for details on this step.

Histogram of data

Fig. 3.6

```
traps <- unique(data[,c("x", "trap")])
# remove 2 NA values
data$sun[is.na(data$sun)] <- mean(data$sun, na.rm=TRUE)
# discretise z into classes
data$z.cut <- cut(data$z, breaks = c(-Inf, 0, 10, 30, Inf))
q <- ggplot(data.frame(table(data$z.cut))) + geom_col(aes(x=Var1, y=Freq)) +
  theme_minimal() + theme(axis.title = element_blank())
q + scale_x_discrete(labels=c("0", "1-10", "11-30", ">30"))
```



Many zeroes!

Principal components analysis

Section 3.4.2.2

A PCA was run on the daily data, the following variables were selected for the PCA:

Response variable:

z (daily count of pollen beetles)

Explanatory variables:

year (The calendar year of the observation) Days_from_start (Number of days from 1st January 1987, the first record in the dataset) day (Julian day number) lat (Latitude) lon (Longitude) rain (Daily rainfall mm) wind (Average windspeed m/s) humidity (Relative humidity %) dew (Daily mean dew point degC) t.min (Minimum Temperature degC) t.max (Maximum temperature degC) sun (Solar radiation) temp (Mean temperature degC) Autumn.Prec (Accumulated rainfall for September, October & Novemer the previous year mm) Winter Prec. (Accumulated rainfall for December the previous year plus January and February mm) Winter.Temp (Accumulated day-degrees above 0 degC for December the previous year plus January and February degC) z.sum (Count of pollen beetles from the trap after June 1st the previous year)

```
#Select data

dailyPCA.unlog<-subset(daily, select = c(year, Days_from_start, day, z, lat,
lon, rain, wind, humidity, dew, t.min, t.max, sun, temp, Autumn.Prec,
Winter.Prec, Winter.Temp, z.sum))

#Remove NAs
dailyPCA.unlog<-dailyPCA.unlog%>% drop_na()
#View data
head(dailyPCA.unlog)

##   year Days_from_start day z      lat       lon rain wind humidity   dew
t.min
## 1 1992            1991 165 0 51.80700 -0.360091 0.01 1.78    62.02  9.83
24.05
## 2 1995            3102 180 5 51.80700 -0.360091 0.01 2.63    64.93 11.23
24.95
## 3 1995            3102 180 2 52.26068  0.568430 0.02 2.37    70.28 11.35
23.29
## 4 1996            3455 168 0 51.80700 -0.360091 0.00 1.82    68.89 10.04
22.78
## 5 1999            4559 176 4 51.80700 -0.360091 0.01 2.31    70.14  9.24
20.30
## 6 2000            4918 170 1 51.80700 -0.360091 0.00 2.84    71.80 14.24
25.66
##   t.max   sun  temp Autumn.Prec Winter.Prec Winter.Temp z.sum
## 1  8.56 31.61 17.21    1.658132   0.8033844      -76.79    47
## 2 10.92 31.57 17.96    2.416264   2.7629989      -33.94   550
## 3 10.11 31.54 16.82    2.124176   2.4285156      -21.36   190
## 4  7.14 31.43 15.84    1.973407   2.3080108     -130.99   234
## 5  8.47 31.21 14.60    2.969560   2.2802351      -50.43    44
## 6 11.95 31.14 19.49    2.039890   2.2130161      -38.82   143

#Run PCA
dailyPCA.unlog.prcomp<-prcomp(dailyPCA.unlog, center=TRUE, scale. = TRUE)

#View
print(dailyPCA.unlog.prcomp)

## Standard deviations (1, .., p=18):
##  [1] 2.3726857150 1.5700840169 1.3107705401 1.2262899236 1.1322440094
```

```
##  [6] 1.0148924483 0.9576248397 0.8865136652 0.8376403125 0.7713377804
## [11] 0.7320672290 0.6531070354 0.4841234340 0.3467649338 0.2196149262
## [16] 0.0779567535 0.0187993893 0.0000517825
## 
## Rotation (n x k) = (18 x 18):
##                          PC1          PC2         PC3         PC4
PC5
## year            0.013832867 -0.584762925 -0.20654404 -0.03111932 -
0.066727955
## Days_from_start  0.021656429 -0.584562200 -0.20519413 -0.03032612 -
0.068232179
## day             0.381976490 -0.010274371  0.05884254  0.03766064 -
0.075073086
## z               0.086864048 -0.052264714 -0.19799190 -0.07725167
0.145658188
## lat            -0.025285769 -0.123276462 -0.13370703  0.04828437 -
0.541564300
## lon             0.012396334  0.349340633 -0.37333799  0.23412610
0.126525834
## rain           -0.044864759 -0.111695729  0.22867363  0.57815212 -
0.078025732
## wind           -0.137091857 -0.049680708  0.07995310  0.37412117
0.074703315
## humidity       -0.286503967 -0.130349796  0.17805679  0.30125085 -
0.063762453
## dew             0.380338896 -0.052007942  0.12937961  0.24010780 -
0.031363044
## t.min           0.412858663  0.026156272  0.03111738  0.04795948 -
0.001113091
## t.max           0.385713963 -0.049539968  0.10366912  0.23279579 -
0.020408811
## sun             0.329812236  0.013333637 -0.10757890 -0.32912561
0.022436990
## temp            0.410834831 -0.008762253  0.06644066  0.13132893 -
0.010219470
## Autumn.Prec    -0.022125916 -0.083566437  0.49728519 -0.22121670 -
0.078399137
## Winter.Prec    -0.004656759 -0.293578722  0.36626827 -0.20082627
0.323552187
## Winter.Temp     0.008849843 -0.113558613  0.01434465  0.11934082
0.704623963
## z.sum           0.012942967 -0.179348624 -0.44590527  0.15929160
0.172243776
##                          PC6          PC7         PC8          PC9
PC10
## year            0.181793391 -0.059127758 -0.02033160 -0.032321880
0.13231087
## Days_from_start  0.181925441 -0.057664790 -0.01965167 -0.032549041
0.12955848
## day             0.012373878  0.069660988  0.03262268 -0.012249151 -
0.13109708
## z              -0.618258541 -0.528675202 -0.48125360  0.105495449
0.06982533
## lat            -0.421606765  0.334092803  0.16633748  0.569265728 -
```

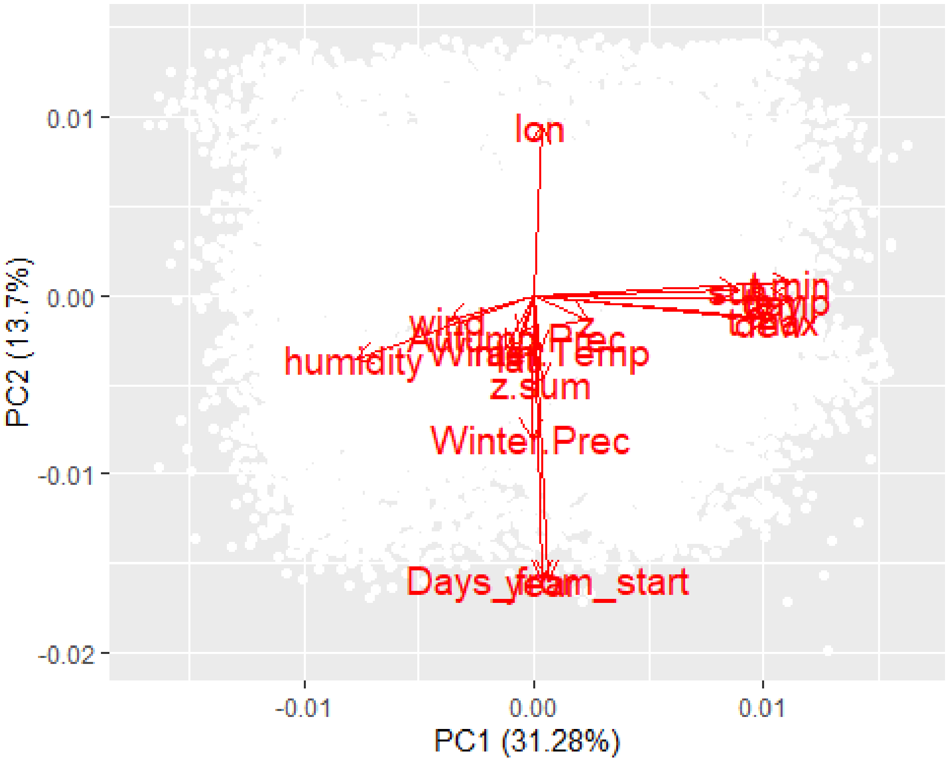
```
0.14867556
## lon             -0.160071578  0.094561545  0.25149744 -0.119074027
0.19003332
## rain            -0.074736130 -0.233458694 -0.04324231 -0.157966862 -
0.56617712
## wind             0.001850209  0.550314417 -0.66613694 -0.018463062
0.21199123
## humidity        -0.077745472 -0.280734392  0.28741735  0.073166938
0.30954546
## dew             -0.014652700 -0.075871357  0.09737765  0.032003361
0.20062020
## t.min            0.024004092 -0.005315734  0.01699053 -0.005323933
0.05287828
## t.max            0.000456389  0.011215425  0.02260155  0.021659348
0.15925841
## sun             -0.015750877  0.113984456 -0.11657524 -0.007833120 -
0.19494927
## temp             0.008683943  0.005138815  0.01046454  0.007859388
0.09387762
## Autumn.Prec     -0.401289894  0.132763385  0.12319523 -0.258732929
0.41797143
## Winter.Prec     -0.239517594  0.204168553  0.11178324 -0.098496760 -
0.35571673
## Winter.Temp      0.002099001  0.087205174  0.18412933  0.556645924
0.04104703
## z.sum           -0.341947138  0.257415864  0.24867172 -0.479147971 -
0.07725741
##                         PC11        PC12         PC13         PC14
## year             0.190641893  0.12867403 -0.021616737 -0.031377728
## Days_from_start  0.192168339  0.12796068 -0.034628432 -0.018307481
## day              0.079612577 -0.02887941 -0.635218199  0.637016441
## z               -0.045674766  0.05738835 -0.091422890  0.014985444
## lat             -0.004670473  0.01767756  0.020002676 -0.078445111
## lon              0.416225572  0.58725963  0.010265423  0.035505582
## rain             0.397854907 -0.09342503  0.112463578 -0.073255511
## wind            -0.039824517  0.06173724  0.062448016  0.141202930
## humidity        -0.275335504  0.06460673  0.290758348  0.525010944
## dew             -0.172695524  0.04509609  0.155155602 -0.017608818
## t.min           -0.062421230  0.01631104  0.122122907 -0.183625525
## t.max           -0.123122100  0.03792030 -0.017720200 -0.177204461
## sun              0.236788132 -0.07386018  0.662240807  0.427832761
## temp            -0.082519684  0.02125531  0.066157450 -0.174473317
## Autumn.Prec      0.449705040 -0.21357658 -0.039958798 -0.069637794
## Winter.Prec     -0.270159706  0.55978311  0.003481400 -0.021812557
## Winter.Temp      0.220088651 -0.27078387 -0.030517698 -0.004140113
## z.sum           -0.271911245 -0.39887433  0.009922098  0.007507915
##                          PC15          PC16          PC17          PC18
## year            -0.0190770633  0.0010255812  0.0006781423 -7.067897e-01
## Days_from_start -0.0201654132  0.0014267591  0.0011681251  7.072751e-01
## day             -0.0538032492  0.0189749151  0.0116537162 -1.449945e-02
## z                0.0004337322  0.0049386846  0.0004885108  3.384415e-07
## lat             -0.0234208987  0.0044474054 -0.0001322247  1.508621e-06
## lon              0.0018427826 -0.0024262386  0.0015570612  3.233934e-06
## rain            -0.0172546053  0.0085109924 -0.0013949272  5.057865e-07
```

```
## wind          -0.0740575252  0.0142991366  0.0021114883  1.091786e-06
## humidity      -0.0548578190  0.1516063891  0.1990259410 -3.101348e-05
## dew           -0.1057066679 -0.4816743052 -0.6489314171  1.199787e-04
## t.min         -0.6285011131  0.6113774602 -0.0168967128 -3.292377e-06
## t.max          0.7398611809  0.4014697453  0.0076617814 -7.526856e-06
## sun            0.1501380798  0.0065591489  0.0060595212  1.375794e-06
## temp          -0.1009154533 -0.4574656022  0.7339945650 -1.279054e-04
## Autumn.Prec   -0.0187488194  0.0049187043  0.0005155531 -2.009229e-05
## Winter.Prec   -0.0019383871 -0.0003658002  0.0001697751  3.003174e-05
## Winter.Temp   -0.0175275857  0.0034716002  0.0004847267 -4.935949e-06
## z.sum         -0.0045729191 -0.0008592697 -0.0007084059  3.263832e-06
```

Plot these values as a biplot

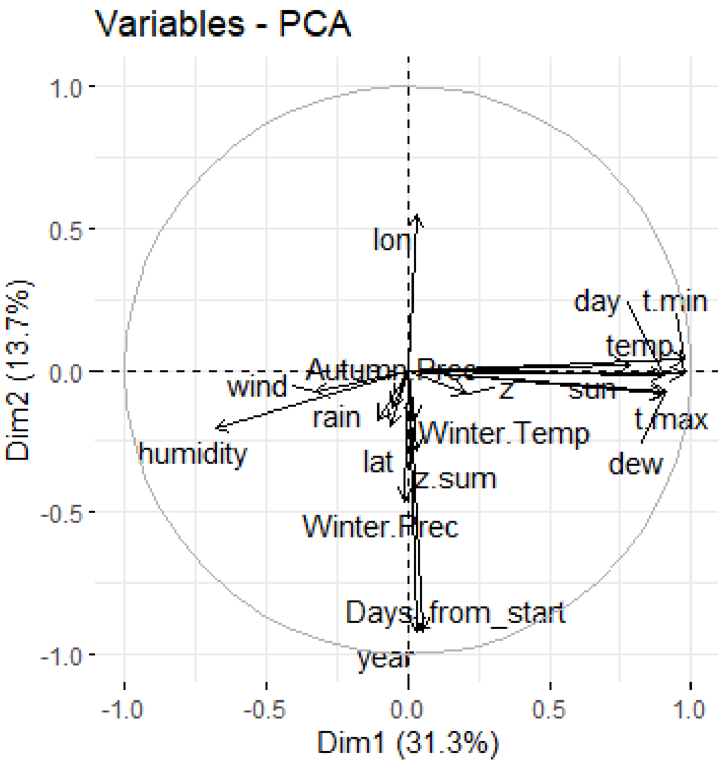
Fig. 3.5

```r
autoplot(dailyPCA.unlog.prcomp, colour="white", loadings = TRUE,
loadings.label=TRUE, loadings.label.size = 5)
```



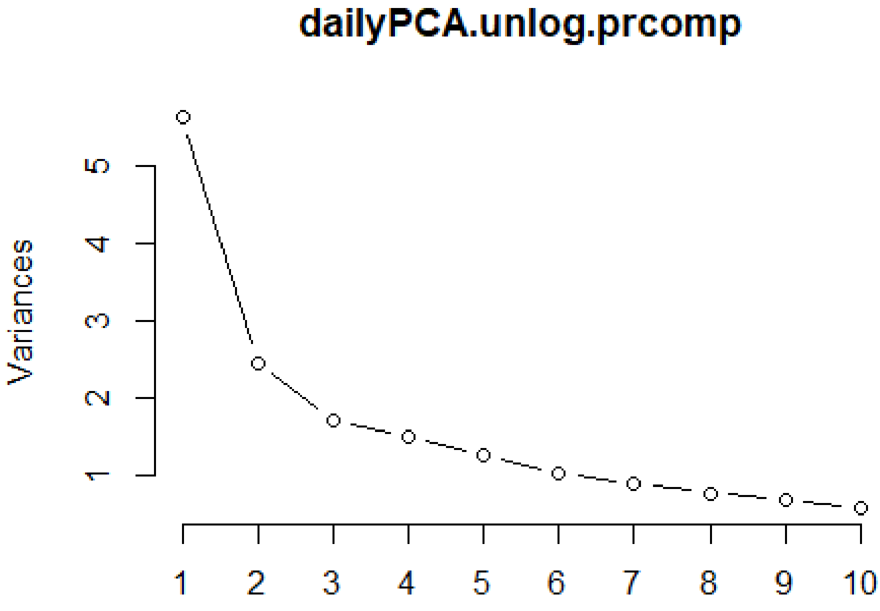
Alternative view

```r
fviz_pca_var(dailyPCA.unlog.prcomp, repel = TRUE)      # Avoid text overlapping
```



Create a Scree plot ## Fig. 3.4

```
plot(dailyPCA.unlog.prcomp, type ="l")
```



Suggests 2-5 PCAs

```
summary(dailyPCA.unlog.prcomp)
```

```
## Importance of components:
##                           PC1    PC2     PC3     PC4     PC5     PC6     PC7
## Standard deviation     2.3727 1.5701 1.31077 1.22629 1.13224 1.01489 0.95762
## Proportion of Variance 0.3128 0.1369 0.09545 0.08354 0.07122 0.05722 0.05095
## Cumulative Proportion  0.3128 0.4497 0.54516 0.62871 0.69993 0.75715 0.80810
##                            PC8     PC9    PC10    PC11   PC12    PC13    PC14
## Standard deviation     0.88651 0.83764 0.77134 0.73207 0.6531 0.48412 0.34676
## Proportion of Variance 0.04366 0.03898 0.03305 0.02977 0.0237 0.01302 0.00668
## Cumulative Proportion  0.85176 0.89074 0.92379 0.95356 0.9773 0.99028 0.99696
##                           PC15    PC16    PC17      PC18
## Standard deviation     0.21961 0.07796 0.01880 5.178e-05
## Proportion of Variance 0.00268 0.00034 0.00002 0.000e+00
## Cumulative Proportion  0.99964 0.99998 1.00000 1.000e+00
```

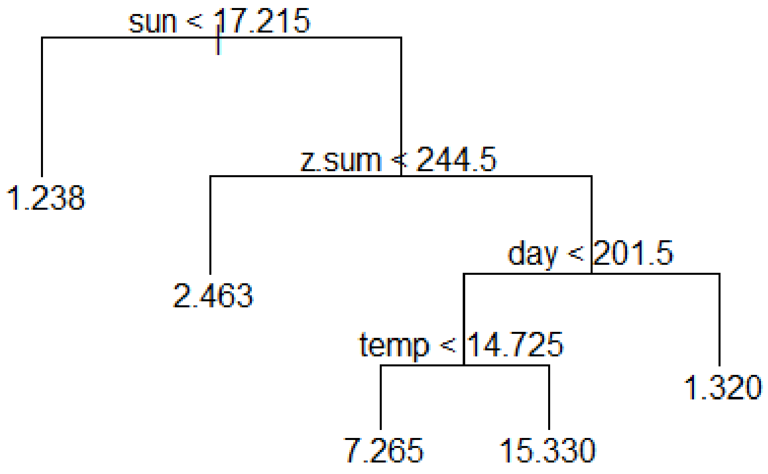
Suggests maybe 10?

Tree model

Section 3.4.2.4

Fig. 3.7

```
dailytree<-subset(daily, select = c(year, trap, Days_from_start, day, z, lat,
lon, rain, wind, humidity, dew, t.min, t.max, sun, temp, Autumn.Prec,
Winter.Prec, Winter.Temp, z.sum))

library(tree)
modelt<-tree(z~.,data=dailytree)
plot(modelt)
text(modelt)
```



Generalized Additive Models

Section 3.4.3

Model selection

Initial model selection looked at the number of variables and various methods of smoothing to find a model that would converge. These steps are available on request, but for the sake of brevity are omitted here.

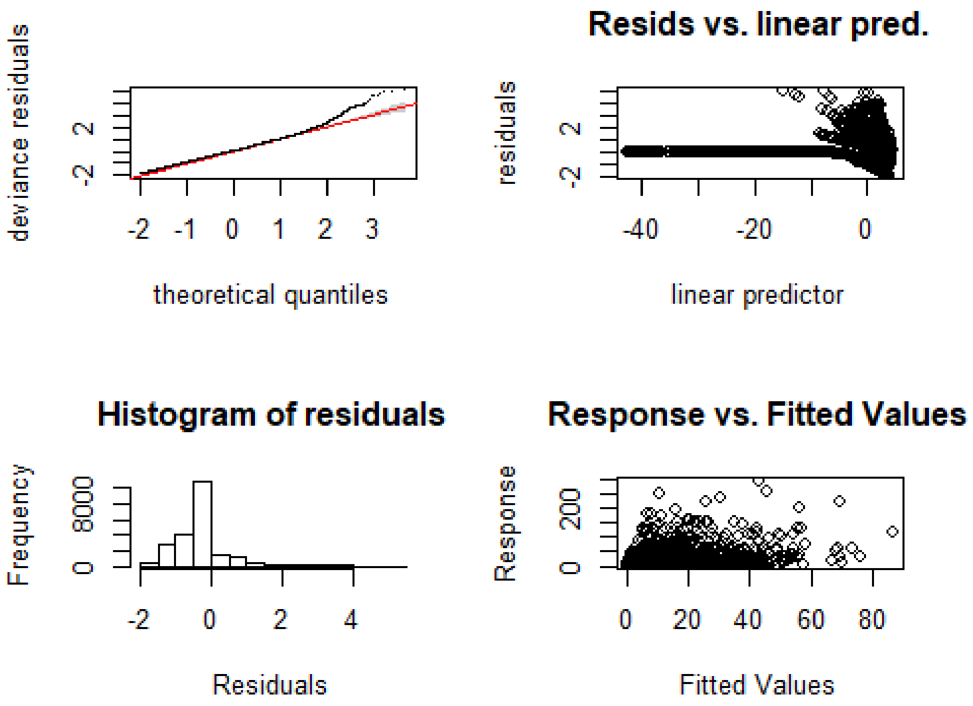
The code for the final model was:

```
gc()

##           used  (Mb) gc trigger  (Mb) max used  (Mb)
## Ncells 4793622 256.1    8602728 459.5  8347882 445.9
## Vcells 9831005  75.1   17829126 136.1 17808555 135.9

Model4sel<-gam(z~s(lon,lat, k=5) +
                 s(Days_from_start, k=100)+s(day, k=100)+
                 s(wind, k=50)+ s(sun,k=100)+ s(temp, k=80)
               + s(Autumn.Prec, k=50)+ s(Winter.Prec, k=50)+ s(Winter.Temp,
k=50)+ s(z.sum, k=50),
               select = TRUE, family = nb(), method = "REML", data=dailytree)
```

Where, “Model” is the final model, “gam” is the term to select the GAM within R, “z” is the response, here the number of B. aeneus per day, the terms preceded by “s(“ are the explanatory variables (see Table 1) with “s” denoting the smooth term for the parameter, “bs=”ts”” denotes the thin plate spline (Wood 2003) used to link the spatial variates (“lat”=latitude, “lon”=longitude), “k” is the knot term which determines the degree of smoothing or wiggliness of the trend, here set for each parameter as the default setting within mgcv (10) was too low, “select=TRUE” allows the model to penalise smooth terms to avoid overfitting, “family=nb()” denotes the negative binomial distribution. The smoothing method is set as Restricted Maximum Likelihood (REML) by method = “REML”, which penalises overly wiggly models (Wood 2017).

Fig. 3.8

```r
gam.check(Model4sel, rep=500)
```



```
##
## Method: REML   Optimizer: outer newton
## full convergence after 13 iterations.
## Gradient range [-0.0002716857,0.0009599165]
## (score 23967.64 & scale 1).
## eigenvalue range [-0.000958787,2705.72].
## Model rank =  626 / 626
##
## Basis dimension (k) checking results. Low p-value (k-index<1) may
## indicate that k is too low, especially if edf is close to k'.
##
##                          k'      edf k-index p-value
## s(lon,lat)          4.00000  0.00261    0.84   0.050 *
## s(Days_from_start) 99.00000 91.93795    0.69  <2e-16 ***
## s(day)             99.00000 36.23367    0.77  <2e-16 ***
## s(wind)            49.00000  4.26959    0.84   0.055 .
## s(sun)             99.00000  7.80364    0.78  <2e-16 ***
## s(temp)            79.00000  7.10865    0.85   0.130
## s(Autumn.Prec)     49.00000 35.41336    0.88   0.910
```

```
## s(Winter.Prec)     49.00000  8.57823    0.88   0.905
## s(Winter.Temp)     49.00000  4.75344    0.88   0.920
## s(z.sum)           49.00000  5.54699    0.86   0.650
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
summary(Model4sel)
```

```
##
## Family: Negative Binomial(0.378)
## Link function: log
##
## Formula:
## z ~ s(lon, lat, k = 5) + s(Days_from_start, k = 100) + s(day,
##     k = 100) + s(wind, k = 50) + s(sun, k = 100) + s(temp, k = 80) +
##     s(Autumn.Prec, k = 50) + s(Winter.Prec, k = 50) + s(Winter.Temp,
##     k = 50) + s(z.sum, k = 50)
##
## Parametric coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -3.7206     0.3115  -11.95   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Approximate significance of smooth terms:
##                          edf Ref.df   Chi.sq p-value
## s(lon,lat)          0.002613      4    0.004 0.00178 **
## s(Days_from_start) 91.937952     99 1889.424 < 2e-16 ***
## s(day)             36.233669     99 1820.945 < 2e-16 ***
## s(wind)             4.269586     49  846.720 < 2e-16 ***
## s(sun)              7.803635     99  147.190 < 2e-16 ***
## s(temp)             7.108652     79  574.281 < 2e-16 ***
## s(Autumn.Prec)     35.413357     49  335.473 < 2e-16 ***
## s(Winter.Prec)      8.578227     49  110.928 < 2e-16 ***
## s(Winter.Temp)      4.753444     49   55.351 < 2e-16 ***
## s(z.sum)            5.546989     49   99.364 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## R-sq.(adj) =  0.257   Deviance explained = 65.9%
## -REML =  23968  Scale est. = 1         n = 21586
```

As the gam.check results show (Fig. 3.8) the GAM does not converge satisfactorily and further work in this direction was deemed to not be useful.

Zero-inflated Generalized Linear Models

Section 3.4.3

Next, a zero-inflated GLM was attempted

```
#Use the "daily" datafile
```

```
#create an all-integer dataframe for correlaton plot
daily1 <-modify_if(daily, is.factor, as.integer)
daily1<-subset(daily1, select = c(year, trap, Days_from_start, day, z, lat, lon,
rain, wind, humidity, dew, t.min, t.max, sun, temp, Autumn.Prec, Winter.Prec,
Winter.Temp, z.sum))
```

The data were split into a training and a testing dataset

```
set.seed(123)
index<-sample(2,nrow(daily),replace = TRUE,p=c(.8,.2))
train<-daily[index==1,]
test<-daily[index==2,]

names(daily)
```

```
##  [1] "year"            "trap"            "date"            "Value.Date"
##  [5] "Days_from_start" "day"             "z"               "lat"
##  [9] "lon"             "x"               "y"               "rain"
## [13] "wind"            "humidity"        "dew"             "t.min"
## [17] "t.max"           "sun"             "temp"            "Autumn.Prec"
## [21] "Winter.Prec"     "Winter.Temp"     "z.sum"
```

```
describe(daily)
```

```
##                 vars     n     mean      sd   median  trimmed     mad
min
## year               1 24399  2007.88    9.66  2012.00  2008.87    7.41
1987.00
## trap*              2 24399     4.93    3.36     6.00     4.63    4.45
1.00
## date*              3 24399  4018.62 2307.05  4022.00  4019.96 2962.23
1.00
## Value.Date         4 24399 39546.05 3544.94 41008.00 39906.95 3033.40
31778.00
## Days_from_start    5 24399  7769.05 3544.94  9231.00  8129.95 3033.40
1.00
## day                6 24399   127.02   71.55   127.00   127.24   91.92
1.00
## z                  7 24399     2.55   11.79     0.00     0.42    0.00
0.00
## lat                8 24399    52.17    0.87    52.12    52.04    0.47
50.63
## lon                9 24399    -0.35    1.07    -0.36    -0.17    1.38     -
3.45
## x                 10 24399   512.80   73.93   513.16   525.15   92.25
297.21
## y                 11 24399   254.45   96.61   247.51   240.36   50.71
82.11
## rain              12 24399     1.91    3.23     0.54     1.16    0.77
0.00
## wind              13 24398     3.47    1.57     3.24     3.33    1.48
0.39
## humidity          14 24399    82.43    8.25    82.17    82.48    9.01
49.07
## dew               15 24399     7.68    4.93     7.86     7.80    5.81    -
```

```
10.42
## t.min             16 24399    14.64    6.43    15.03    14.71    7.69    -
7.13
## t.max             17 24399     6.76    5.03     6.86     6.86    6.05    -
8.84
## sun               18 24393    12.61    7.29    12.38    12.33    8.82
0.07
## temp              19 24399    10.67    5.72    10.86    10.79    7.00    -
8.13
## Autumn.Prec       20 24399     2.11    0.50     2.06     2.09    0.44
1.11
## Winter.Prec       21 24399     2.08    0.69     2.10     2.03    0.74
0.80
## Winter.Temp       22 24399   -49.49   47.32   -42.76   -43.75   31.73  -
401.23
## z.sum             23 21591   396.81  445.72   210.00   305.38  210.53
0.00
##                      max    range  skew kurtosis    se
## year             2018.00    31.00 -0.66    -1.00  0.06
## trap*              12.00    11.00  0.33    -0.87  0.02
## date*            8000.00  7999.00  0.00    -1.20 14.77
## Value.Date      43350.00 11572.00 -0.65    -1.01 22.69
## Days_from_start 11573.00 11572.00 -0.65    -1.01 22.69
## day              250.00   249.00 -0.02    -1.18  0.46
## z                593.00   593.00 15.17   439.55  0.08
## lat               55.21     4.58  1.82     4.29  0.01
## lon                0.94     4.40 -1.32     1.27  0.01
## x                605.90   308.69 -1.33     1.32  0.47
## y                591.05   508.94  1.80     4.24  0.62
## rain              46.23    46.23  3.37    17.47  0.02
## wind              12.17    11.78  0.91     1.09  0.01
## humidity         100.00    50.93 -0.07    -0.53  0.05
## dew               19.48    29.90 -0.21    -0.71  0.03
## t.min             34.23    41.36 -0.09    -0.87  0.04
## t.max             19.23    28.07 -0.13    -0.89  0.03
## sun               31.61    31.54  0.23    -0.92  0.05
## temp              25.18    33.31 -0.14    -0.94  0.04
## Autumn.Prec        4.14     3.04  0.90     2.23  0.00
## Winter.Prec        4.45     3.65  0.77     0.93  0.00
## Winter.Temp        0.00   401.23 -3.73    24.19  0.30
## z.sum            2286.00  2286.00  2.00     4.22  3.03
```

The data were then centred. This prevents a true forecast as the data would have to be rescaled back to the original values post-processing.

```
year.scaled<-scale(daily$year, scale=TRUE)
Days_from_start.scaled<-scale(daily$Days_from_start, scale=TRUE)
day.scaled<-scale(daily$day, scale=TRUE)
daily.beetle.count.scaled<-scale(daily$z, scale=TRUE)
lat.scaled<-scale(daily$lat, scale=TRUE)
lon.scaled<-scale(daily$lon, scale=TRUE)
rain.scaled<-scale(daily$rain, scale=TRUE)
wind.scaled<-scale(daily$wind, scale=TRUE)
humidity.scaled<-scale(daily$humidity, scale=TRUE)
```

```
dew.scaled<-scale(daily$dew, scale=TRUE)
t.min.scaled<-scale(daily$t.min, scale=TRUE)
t.max.scaled<-scale(daily$t.max, scale=TRUE)
sun.scaled<-scale(daily$sun, scale=TRUE)
temp.scaled<-scale(daily$temp, scale=TRUE)
Autumn.Prec.scaled<-scale(daily$Autumn.Prec, scale=TRUE)
Winter.Prec.scaled<-scale(daily$Winter.Prec, scale=TRUE)
Winter.Temp.scaled<-scale(daily$Winter.Temp, scale=TRUE)
previous.pbcount.scaled<-scale(daily$z.sum, scale=TRUE)
```

As the daily data for the pollen beetles contained fractions where multi-day catches cought a number of beetles not wholly divisible from the number of days trapping these data had to be rounded and then scaled.

```
rounded.beetle.count<-round(daily$z, digits=0)
scaled.rounded.beetle.count<-scale(rounded.beetle.count, scale=TRUE)
```

A process of model selection was undergone, details of which are excluded from this summary but the full code is available on request.

The final model code in R was:

```
mod13sel<-zeroinfl(rounded.beetle.count  ~ previous.pbcount.scaled + temp.scaled
+ wind.scaled + lon.scaled |sun.scaled + wind.scaled + day.scaled + lon.scaled,
dist = "negbin")

summary(mod13sel)

##
## Call:
## zeroinfl(formula = rounded.beetle.count ~ previous.pbcount.scaled +
temp.scaled +
##     wind.scaled + lon.scaled | sun.scaled + wind.scaled + day.scaled +
##     lon.scaled, dist = "negbin")
##
## Pearson residuals:
##     Min      1Q  Median      3Q     Max
## -0.4993 -0.3651 -0.2320 -0.1385 35.3899
##
## Count model coefficients (negbin with log link):
##                         Estimate Std. Error z value Pr(>|z|)
## (Intercept)              0.78571    0.03211   24.47  < 2e-16 ***
## previous.pbcount.scaled  0.52684    0.02225   23.68  < 2e-16 ***
## temp.scaled              0.70084    0.03582   19.57  < 2e-16 ***
## wind.scaled             -0.52678    0.02951  -17.85  < 2e-16 ***
## lon.scaled              -0.08125    0.02500   -3.25  0.00115 **
## Log(theta)              -1.37363    0.03428  -40.07  < 2e-16 ***
##
## Zero-inflation model coefficients (binomial with logit link):
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -0.40973    0.08904  -4.602 4.19e-06 ***
## sun.scaled  -2.13522    0.10518 -20.300  < 2e-16 ***
## wind.scaled  0.23727    0.04598   5.160 2.47e-07 ***
## day.scaled   0.67251    0.05206  12.918  < 2e-16 ***
## lon.scaled  -0.21479    0.04227  -5.081 3.76e-07 ***
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Theta = 0.2532
## Number of iterations in BFGS optimization: 24
## Log-likelihood: -2.646e+04 on 11 Df
```

The vertical bar ‘|’ defines the model partition in terms of where the parameters are utilised. Those predictors on the left side being used for the count (negative binomial) component and those predictors on the right the zero-inflation (binomial) component of the model. In other words, those on the right side are explaining whether or not there is a risk of a zero count, with those on the left determining the size of the count when it is non-zero.

For predictive purposes the NA values were removed

```
df<-cbind(rounded.beetle.count, daily$wind, daily$sun, daily$z.sum)
df1<-na.exclude(df)
```

The observed values were stored and the number of nonzero values determined:

```
obs <- df1[,1]
obs.nonzero <- obs > 0
table(obs)

## obs
##     0     1     2     3     4     5     6     7     8     9    10    11    12
## 15852  1635   879   537   389   282   212   198   155   135    94    96    80
##    13    14    15    16    17    18    19    20    21    22    23    24    25
##    83    74    48    43    38    35    43    38    33    41    27    19    18
##    26    27    28    29    30    31    32    33    34    35    36    37    38
##    24    23    17    21    18    14    12    19    17    18    18    20    13
##    39    40    41    42    43    44    45    46    47    48    49    50    51
##     8     9     4     4     7     4    11     7     5     7     7     8     6
##    52    53    54    55    56    57    58    59    60    61    62    63    64
##    10    11     5     2     7     3     4     2     7     2     9     4     3
##    65    66    67    68    69    72    73    74    75    76    77    79    80
##     4     2     2     1     2     2     3     3     1     2     4     3     1
##    81    82    83    85    86    88    90    91    92    93    94    95    96
##     1     1     4     2     3     2     1     1     1     1     4     1     4
##    97    99   100   103   105   106   107   108   109   111   116   120   124
##     1     1     3     1     1     2     2     1     1     1     3     1     1
##   125   127   128   129   131   132   135   136   141   148   150   151   152
##     1     3     4     1     1     3     1     1     2     1     1     1     2
##   154   157   159   174   179   196   200   219   222   233   248   257   295
##     2     1     1     1     2     1     1     1     1     1     1     1     1

table(obs.nonzero)

## obs.nonzero
## FALSE  TRUE
## 15852  5734
```

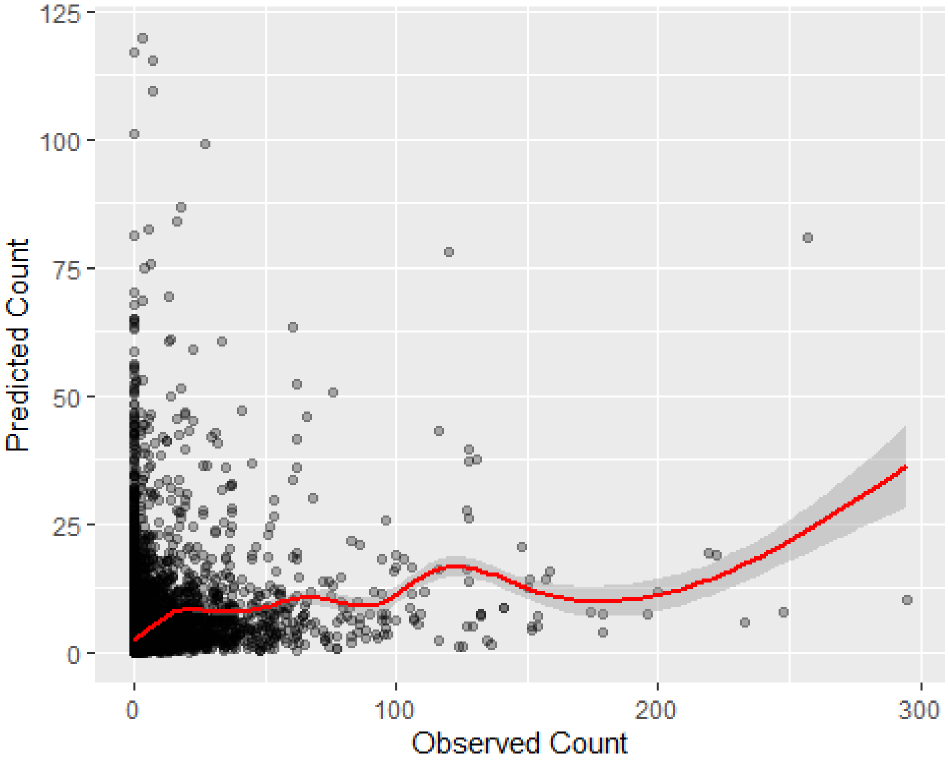
The predicted counts were stored and the number of non-zero counts determined:

```
preds.count <- predict(mod13sel, type="response")

preds <- 1-predict(mod13sel, type = "prob")[,1]
preds.nonzero <- preds > 0.5
table(preds.nonzero)

## preds.nonzero
## FALSE  TRUE
## 18708  2878
```

The observed versus predicted could then be calculated:

Fig. 3.10

```
output <- as.data.frame(list(preds.count=preds.count, obs=obs))
#output2 <- stri_list2matrix(output)
summary(output)

##   preds.count             obs
##  Min.   :  0.00212   Min.   :  0.000
##  1st Qu.:  0.29696   1st Qu.:  0.000
##  Median :  1.16003   Median :  0.000
##  Mean   :  2.85777   Mean   :  2.471
##  3rd Qu.:  3.25244   3rd Qu.:  1.000
##  Max.   :119.76060   Max.   :295.000

#class(output2) <- 'numeric'

r<- ggplot(aes(x=obs, y=preds.count), data=output) + geom_point(alpha=0.3) +
geom_smooth(col="red")
r+ labs(x = "Observed Count", y = "Predicted Count")

## `geom_smooth()` using method = 'gam' and formula 'y ~ s(x, bs = "cs")'
```



Forecasting using the model.

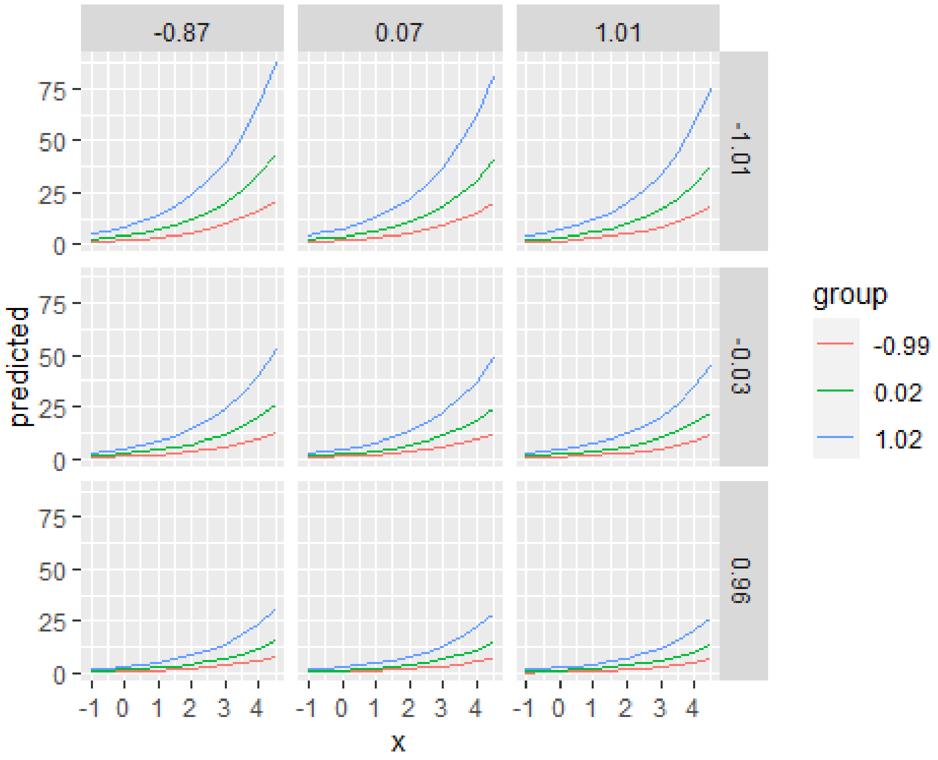
Taking third value from each scaled column plus a range within the range of scaled day a forecast can be made

```
newdata <- expand.grid(list(previous.pbcount.scaled=3.103,
                            temp.scaled=-2.006,
                            wind.scaled=-0.941,
                            lon.scaled=-0.0061,
                            sun.scaled=-1.29,
                            day.scaled=-1.7:1.7))

predict(mod13sel, newdata, se.fit = FALSE, conf = 0.95,
        MC = 1000, type = "response",
        na.action = na.pass)

##         1         2         3         4
## 1.2367754 0.7286707 0.4037233 0.2154716

r2_zeroinflated(mod13sel)

## # R2 for Zero-Inflated and Hurdle Regression
##
##        R2: 0.953
##   adj. R2: 0.953

coef(mod13sel)

##              count_(Intercept) count_previous.pbcount.scaled
##                     0.78571422                    0.52683651
##              count_temp.scaled             count_wind.scaled
##                     0.70083707                   -0.52678207
```

```
##               count_lon.scaled             zero_(Intercept)
##                    -0.08125415                  -0.40973257
##                zero_sun.scaled              zero_wind.scaled
##                    -2.13521825                   0.23726687
##                zero_day.scaled               zero_lon.scaled
##                     0.67250901                  -0.21478931
```

...and these forecasts plotted

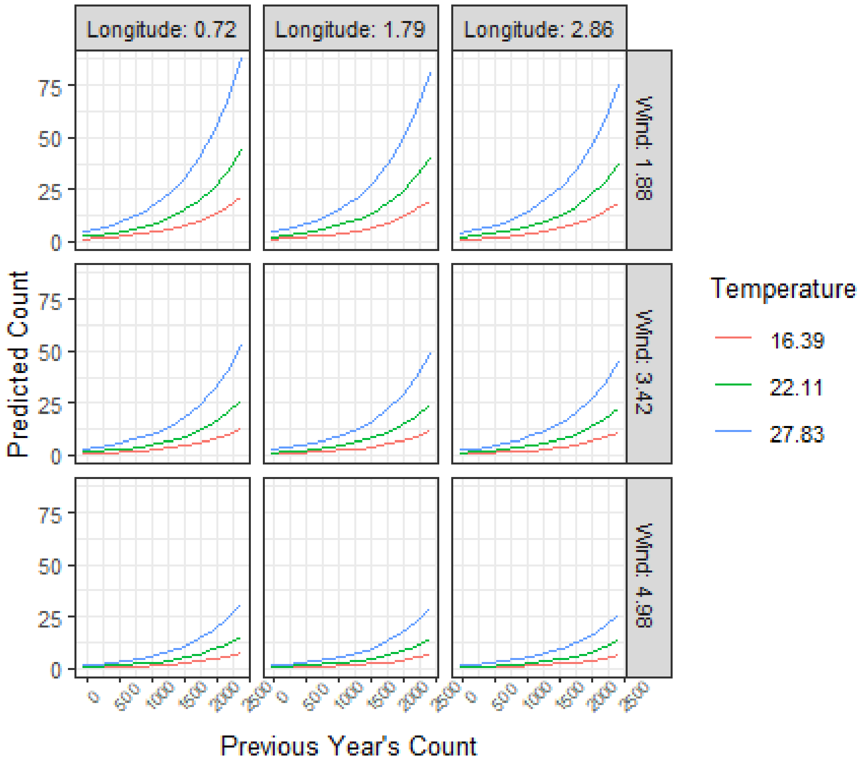
```
ggmodprev.temp.lon.wind<-ggpredict(mod13sel, terms =
c("previous.pbcount.scaled", "temp.scaled", "lon.scaled", "wind.scaled"))

ggplot(ggmodprev.temp.lon.wind,aes(x, predicted, colour = group)) +
  geom_line() +
  facet_grid(panel~facet)
```



Then unscale the variables and rerun ## Fig. 3.9

```
#turn the ggpredict output into a dataframe
labelled.ggmod<-as_tibble(ggmodprev.temp.lon.wind)
#relabel the columns
labelled.ggmod<-labelled.ggmod %>%
  rename(
    Previous = x,
    Temperature = group,
    Longitude = facet,
    Wind = panel)

#make a separate dataframe for unscaling
labelled.ggmod2<-labelled.ggmod
```

```r
#convert the columns from factors to numeric
labelled.ggmod2$Temperature <- as.numeric(labelled.ggmod2$Temperature)
labelled.ggmod2$Wind <- as.numeric(labelled.ggmod2$Wind)
labelled.ggmod2$Longitude <- as.numeric(labelled.ggmod2$Longitude)
labelled.ggmod2$Previous <- as.numeric(labelled.ggmod2$Previous)

#remove the NAs for the mean on Previous & wind
wind.noNA<-na.exclude(daily$wind)
Prev.noNA<-na.exclude(daily$z.sum)
#unscale
labelled.ggmod2$Temperature <- (labelled.ggmod2$Temperature * sd(daily$temp)) +
mean(daily$temp)
labelled.ggmod2$Wind <- (labelled.ggmod2$Wind * sd(wind.noNA)) + mean(wind.noNA)
labelled.ggmod2$Longitude <- (labelled.ggmod2$Longitude * sd(daily$lon)) +
mean(daily$lon)
labelled.ggmod2$Previous <- (labelled.ggmod2$Previous * sd(Prev.noNA)) +
mean(Prev.noNA)
#round
labelled.ggmod2$Temperature <- round(labelled.ggmod2$Temperature, digits=2)
labelled.ggmod2$Wind <- round(labelled.ggmod2$Wind, digits=2)
labelled.ggmod2$Longitude <- round(labelled.ggmod2$Longitude, digits=2)
labelled.ggmod2$Previous <- round(labelled.ggmod2$Previous, digits=2)
write.csv(labelled.ggmod2,"P://Analysis/Zero Inflated Models/ggmod2.csv",
row.names = FALSE)



p<-ggplot(labelled.ggmod2,aes(Previous, predicted, colour =
factor(Temperature))) +
  geom_line() +
  facet_grid(Wind~Longitude, labeller=label_both)

p<- p + labs(colour="Temperature")
p<-p + labs(x = "Previous Year's Count", y = "Predicted
Count")+theme_bw(base_size=10)
p<-p + theme(axis.text.x = element_text(size=6, angle=45))
p + coord_fixed(ratio = 32)
```



Random Forest

Section 3.4.4.1

To investigate the Random Forest method for understanding the data the packages randomForest, Rpart and Rpartykit were used.

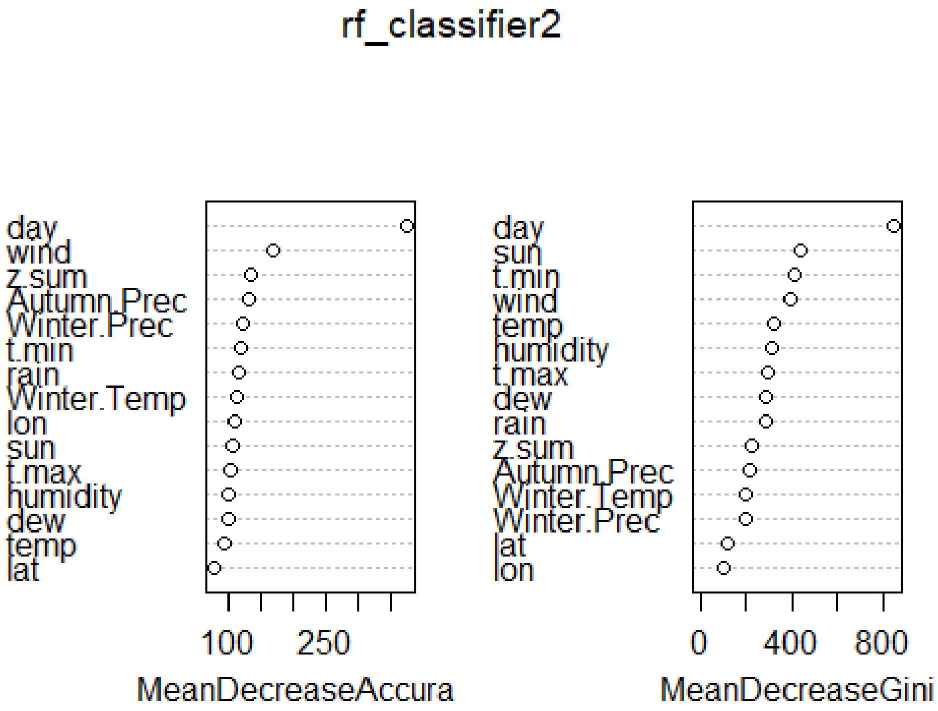
First, the dataset was assembled

```
#Use the daily datafile
# 2 NA values
daily$sun[is.na(daily$sun)] <- mean(daily$sun, na.rm=TRUE)
daily1 <-modify_if(daily, is.factor, as.integer)
daily1<-na.omit(daily1)

data2<-subset(daily1, select = c(day, lat, lon, rain, wind, humidity, dew,
t.min, t.max, sun, temp, Autumn.Prec, Winter.Prec, Winter.Temp, z.sum))

data2$z.cut <- cut(daily1$z, breaks = c(-Inf, 0, 10, 30, Inf))
```

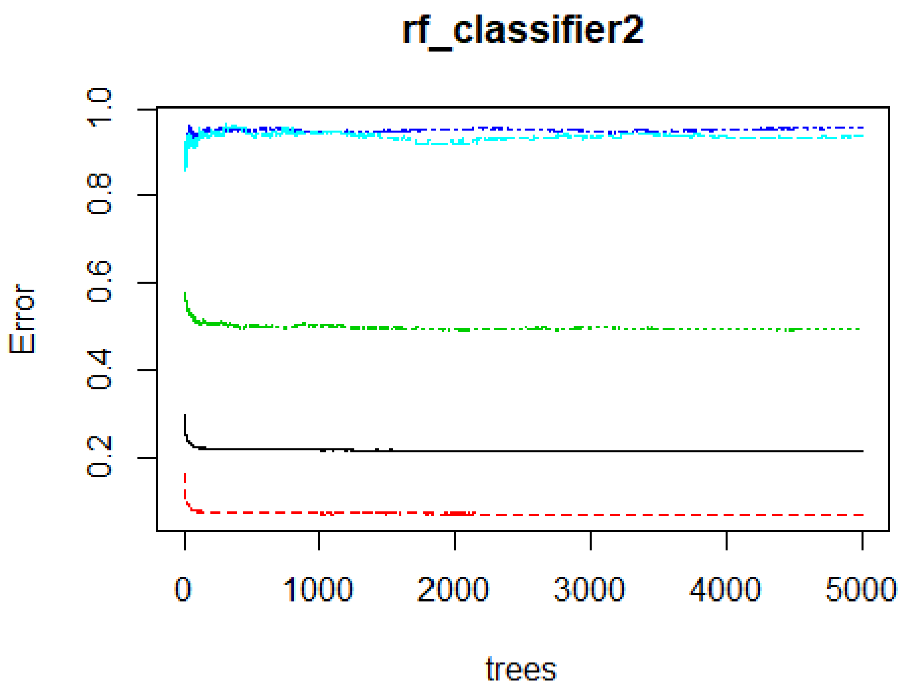
The random forest model was created

```
# Set random seed to make results reproducible:
set.seed(17)
# Calculate the size of each of the data sets:
data_set_size2 <- floor(nrow(data2)/2)
# Generate a random sample of "data_set_size" indexes
indexes2 <- sample(1:nrow(data2), size = data_set_size2)
# Assign the data to the correct sets
training2 <- data2[indexes2,]
```

```
validation2 <- data2[-indexes2,]
rf_classifier2 = randomForest(z.cut ~ ., data=training2, ntree=5000, mtry=4,
importance=TRUE)
#ntree is the number of trees, mtry is number of variables to try (4 is max
here)
rf_classifier2

##
## Call:
##  randomForest(formula = z.cut ~ ., data = training2, ntree = 5000,      mtry
= 4, importance = TRUE)
##                Type of random forest: classification
##                      Number of trees: 5000
## No. of variables tried at each split: 4
##
##         OOB estimate of  error rate: 21.42%
## Confusion matrix:
##            (-Inf,0] (0,10] (10,30] (30, Inf] class.error
## (-Inf,0]       7228    536       8         3   0.0703537
## (0,10]         1180   1225      14         1   0.4938017
## (10,30]         127    275      19         3   0.9551887
## (30, Inf]        55    103       7        11   0.9375000

varImpPlot(rf_classifier2)
```

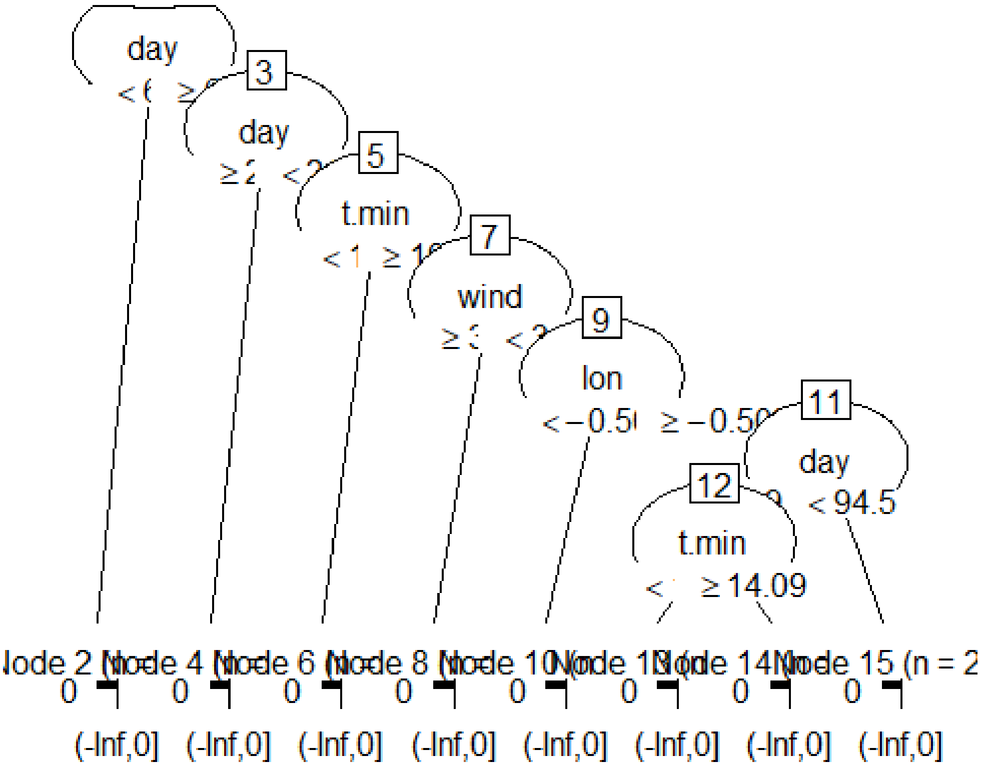


```
plot(rf_classifier2)
```



##Fig. 3.12

```
fit2<-rpart(z.cut ~ ., data=training2)
plot(as.party(fit2), col= cols)
```



To determine the Gini index for each variable the following code was run

```
first_seed <- 50
accuracies <-c()
Importance<-c()
result<-c()
gc()

##            used  (Mb) gc trigger   (Mb)  max used   (Mb)
## Ncells  5052812 269.9    8602728  459.5   8602728  459.5
## Vcells 72445526 552.8  589099977 4494.5 689780662 5262.7

#for (i in 1:10){
#  set.seed(first_seed)
#  first_seed <- first_seed+1
#  modelFit <- randomForest(z.cut ~ ., #data=training2, ntree=5000, proximity =
TRUE,importance =TRUE, votes=TRUE, mytry = 4)
#  print (modelFit)
#  result[[i]] <- modelFit
#  write.csv(modelFit$importance, file = #paste("Importance_", i, ".csv", sep =
""))
#}
#Getting a memory error message after 7 iterations slight change for the last 3
```

These values were written to csv files and then merged to create a single dataframe. Note, this assembly takes a few hours of PC processing time.

```
#Imp1 <- read.csv("Importance_1.csv")
#Imp2 <- read.csv("Importance_2.csv")
#
#...
#
#Imp9 <- read.csv("Importance2_2.csv")
#Imp10 <- read.csv("Importance2_3.csv")

#total <- merge(Imp1 ,Imp2 ,by="X")
#Total <-merge(total ,Imp3 ,by="X")
#
#...

#Total <-merge(Total ,Imp9 ,by="X")
#Total <-merge(Total ,Imp10 ,by="X")
#names (Total)

#Assembled dataset for loading in markdown
Total<-read.csv("P://Final datasets/RFTotal.csv")
```

Calculate the mean, standard error and standard deviation and re-order by the mean:

```
names (Total)

##  [1] "X"                         "X..Inf.0..x"
##  [3] "X.0.10..x"                 "X.10.30..x"
##  [5] "X.30.100..x"               "X.100..Inf..x"
##  [7] "MeanDecreaseAccuracy.x"    "MeanDecreaseGini.x"
##  [9] "X..Inf.0..y"               "X.0.10..y"
## [11] "X.10.30..y"                "X.30.100..y"
```

```
## [13] "X.100..Inf..y"              "MeanDecreaseAccuracy.y"
## [15] "MeanDecreaseGini.y"         "X..Inf.0..x.1"
## [17] "X.0.10..x.1"                "X.10.30..x.1"
## [19] "X.30.100..x.1"              "X.100..Inf..x.1"
## [21] "MeanDecreaseAccuracy.x.1" "MeanDecreaseGini.x.1"
## [23] "X..Inf.0..y.1"              "X.0.10..y.1"
## [25] "X.10.30..y.1"               "X.30.100..y.1"
## [27] "X.100..Inf..y.1"            "MeanDecreaseAccuracy.y.1"
## [29] "MeanDecreaseGini.y.1"       "X..Inf.0..x.2"
## [31] "X.0.10..x.2"                "X.10.30..x.2"
## [33] "X.30.100..x.2"              "X.100..Inf..x.2"
## [35] "MeanDecreaseAccuracy.x.2" "MeanDecreaseGini.x.2"
## [37] "X..Inf.0..y.2"              "X.0.10..y.2"
## [39] "X.10.30..y.2"               "X.30.100..y.2"
## [41] "X.100..Inf..y.2"            "MeanDecreaseAccuracy.y.2"
## [43] "MeanDecreaseGini.y.2"       "X..Inf.0..x.3"
## [45] "X.0.10..x.3"                "X.10.30..x.3"
## [47] "X.30.100."                  "X.100..Inf."
## [49] "MeanDecreaseAccuracy.x.3" "MeanDecreaseGini.x.3"
## [51] "X..Inf.0..y.3"              "X.0.10..y.3"
## [53] "X.10.30..y.3"               "X.30..Inf..x"
## [55] "MeanDecreaseAccuracy.y.3" "MeanDecreaseGini.y.3"
## [57] "X..Inf.0..x.4"              "X.0.10..x.4"
## [59] "X.10.30..x.4"               "X.30..Inf..y"
## [61] "MeanDecreaseAccuracy.x.4" "MeanDecreaseGini.x.4"
## [63] "X..Inf.0..y.4"              "X.0.10..y.4"
## [65] "X.10.30..y.4"               "X.30..Inf."
## [67] "MeanDecreaseAccuracy.y.4" "MeanDecreaseGini.y.4"
```

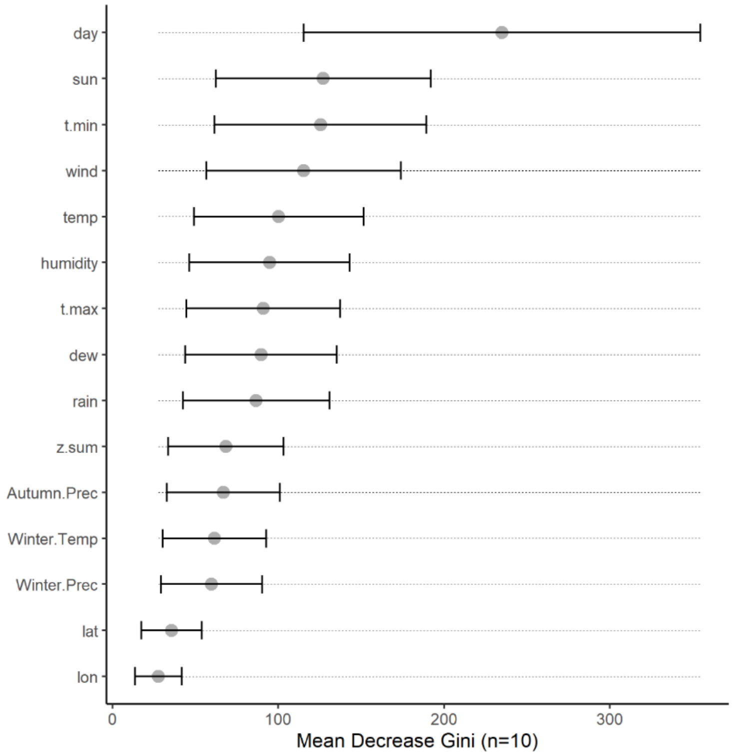
```
Total.Plot<-Total[,c(1,5,9,13,17,21,25,29,33,37,41)]
Total.Plot$Mean<-rowMeans (Total.Plot[,2:11])
Total.Plot$SE<- apply(Total.Plot[,2:11],1,std.error)
Total.Plot$SD<- apply(Total.Plot[,2:11],1,sd)
Total.Plot <- Total.Plot[order(-Total.Plot$Mean),]
```

Make the plot:

Fig. 3.11 Fig. 3.14

```
p<-ggplot(Total.Plot, aes(reorder(X, Mean),Mean)) +
  geom_point(col="#AEAEAE", size=3) + # Draw points
  geom_errorbar(aes(ymin=Mean-SD, ymax=Mean+SD), width=.4,
              position=position_dodge(.9)) +
  geom_segment(aes(x=X,
                   xend=X,
                   y=min(Mean),
                   yend=max(Mean)+max(SD)),
               linetype="dashed",
               size=0.1) +   # Draw dashed lines
  ylab("Mean Decrease Gini (n=10)")+
  xlab("")+
  coord_flip()+
```

```
  theme_bw()+
  theme(axis.line = element_line(colour = "black"),
        panel.grid.major = element_blank(),
        panel.grid.minor = element_blank(),
        panel.border = element_blank(),
        panel.background = element_blank(),
        legend.position="none",
```



```
        legend.title=element_blank())

p
```

Artificial Neural Networks

Section 3.4.5

Note: the following code requires tensorflow to be installed

An ordinal regression artificial neural network was employed as a final attempt to model the daily data.

First, assemble the dataset:

```
traps <- unique(data[,c("x", "trap")])
# remove 2 NA values
data$sun[is.na(data$sun)] <- mean(data$sun, na.rm=TRUE)


# discretise z into classes
data$z.cut <- cut(data$z, breaks = c(-Inf, 0, 10, 30, Inf))
```

```
data <- data[,c("day", "year", "x", "y", "z", "z.cut", "rain", "wind",
"humidity",
                "dew", "t.min", "t.max", "sun", "temp", "z.sum")]

data <- data[order(data$year, data$day, data$x, data$y),]



# normalise the numeric variables
sc <- cbind(rbind(apply(data[,-6], 2, mean), apply(data[,-6], 2, sd)),
z.cut=c(NA, NA))
sc <- sc[, colnames(data)]
```

Keras works with what are know as long short-term memory (LSTM) layers. These need to be created (“shaped”) using the following function:

```
lstm.shape <- function(table, # data.frame of covariates X and response var. Y
                       by=c("x", "y"), # dimension by which to group the
data.frame
                       t, # time dimension along which to wrap the table
                       Y="z", # reponse variable
                       step.X=(-5):(-1), # time step relative to step.Y to use
for covariates
                       step.Y=0, # time step of Y to predict
                       use.Y=TRUE, # include previous value of response var. as
predictors
                       scales=NULL, # a 2-line (mean, sd) matrix by which X and
Y are to be normalised
                       last=FALSE){ # include only the part of X that will
predict the last steps of Y

  arr.X <- array(NA, dim=c(nrow(table), length(step.X), ncol(table)))
  arr.Y <- array(NA, dim=c(nrow(table), length(step.Y), 1))

  t.Y <- sapply(step.Y, function(i) table[, t]+i)
  t.X <- sapply(step.X, function(i) table[, t]+i)

  for (i in 1:length(step.Y)){
    small.table <- cbind(t.Y[,i], table[, by])
    names(small.table)[1] <- t

    temp <- left_join(small.table, table, by=c(t, by))[,Y]
    if (sum(is.na(scales[,Y])) == 0) {temp <- (temp - scales[1,Y]) /
scales[2,Y]}
    arr.Y[,i,] <- temp
  }

  for (i in 1:length(step.X)){
    small.table <- cbind(t.X[,i], table[, by])
    names(small.table)[1] <- t
    temp <- left_join(small.table, table, by=c(t, by))
```

```
    for (j in 1:ncol(table)){
      if (sum(is.na(scales[,names(table)[j]])) == 0) {
        arr.X[,i,j] <- (temp[,j] - scales[1, names(table)[j]]) / scales[2,
names(table)[j]]
      }else{
        arr.X[,i,j] <- temp[,j]
      }
    }
  }

  idx <- which(names(table) == Y)
  if (!use.Y){ # use Y as a predictor; only sensible if !(step.Y %in% step.X)
    arr.X <- arr.X[,,-idx, drop=FALSE]}

  nas <- is.na(arr.X[,,(1 + length(t) + length(by)), drop=FALSE])
  weights <- 1 / (apply(nas, 1, sum) + 1)
  arr.X[is.na(arr.X)] <- 0

  if (last){
    arr.X <- arr.X[t.Y==max(t.Y),,,drop=FALSE]
    arr.Y <- arr.Y[t.Y==max(t.Y),,,drop=FALSE]
    weights <- weights[t.Y==max(t.Y)]
  }

  list(arr.X, arr.Y, weights)
}
```

The architecture of the Neural Network is then created:

```
## number of LSTM/GRU memory time steps
STEPS <-20

## use covariates until today to predict today's cut counting
arr <- lstm.shape(table=data[, names(data)!="z"],
                  by=c("year", "x", "y"), t="day", Y="z.cut",
                  step.X=(-STEPS):0, step.Y=0, use.Y=FALSE,
                  scales=sc[, names(data)!="z"])

## separate labels from data, and drop the first STEPS days
train_data <- arr[[1]][data$day > STEPS,,, drop=FALSE] #[,,-(1:4), drop=FALSE] #
rm coords?
train_labels <- labels_1d <- arr[[2]][data$day > STEPS,,, drop=FALSE]
```

```
samp_weights <- arr[[3]][data$day > STEPS]
## also drop the first STEPS days from original data set
data <- filter(data, day>STEPS)


## this would become problematic if step.Y!=0 like in e.g. X1,X2,X3 -> Y1,Y2,Y3
train_labels <- array(train_labels-1, dim=c(dim(train_labels)[1], 1))
## train_labels <- to_categorical(train_labels) # if one_hot CCE

## ordinal regression one-hot
table <- matrix(0, nrow(train_labels), length(unique(train_labels)))
for(i in 1:nrow(train_labels)){ table[i, 1:(train_labels[i]+1)] <- 1 }
train_labels <- table[,-1]

## training/testing set, structured
who_test <- which(data$year %in% 1989:2004 & data$x==traps$x[traps$trap=='Brooms Barn'])
## who_test <- sort((1:nrow(train_data))[sample(1:nrow(train_data), floor(.3*nrow(train_data)))])
who_train <- (1:nrow(train_data))[!(1:nrow(train_data)) %in% who_test]

## class weights
class_weights <- 1/sqrt(table(data$z.cut[who_train])) * sqrt(nrow(data[who_train,]))
names(class_weights) <- as.character(0:(length(table(data$z.cut))-1))
## aggregate class weights and sample weights in one single vector (required by keras)
samp_weights <- samp_weights * class_weights[labels_1d]

## set seed
use_session_with_seed(42)
```

```
## network architecture
model <- keras_model_sequential()

model %>%
  layer_gru(units = 128, input_shape = dim(train_data)[2:3],
            return_sequences = FALSE, recurrent_dropout = .2,
            kernel_regularizer = regularizer_l2(.001)) %>%
  layer_dense(units = max(max(train_labels)+1, ncol(train_labels)),
              activation ='sigmoid')

# summary(model)
model %>% compile(
  optimizer = optimizer_adam(lr=.0002),
  loss = 'binary_crossentropy'
)
```

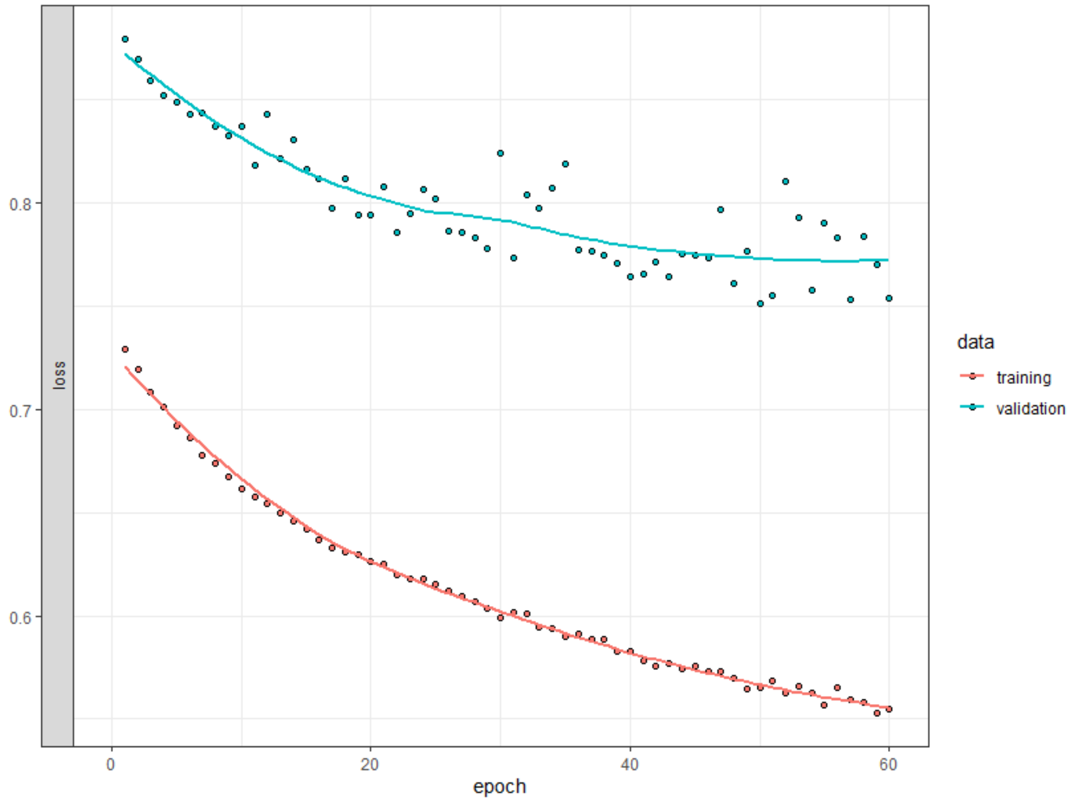
The network is then trained on the dataset:

```
history <- model %>% keras::fit(
  x=train_data[who_train,,, drop=FALSE],
  y=train_labels[who_train,, drop=FALSE],
  validation_split=.2,
  epochs = 100,
  batch_size = 10,
  sample_weight = samp_weights[who_train],
  callbacks = callback_early_stopping(
    monitor="val_loss", patience = 10, restore_best_weights = TRUE),
  shuffle=TRUE
)
```

Fig. 3.14

```
(p2 <- plot(history) + theme_bw() + labs(y="") +
```

```
  scale_x_continuous(limits=c(0, length(history$metrics$val_loss))))
```



```
## save model weights

model %>% save_model_weights_hdf5("D:/pollenBeetle/Final/weightsFinal.hdf5")
```

We then test the model predictions:

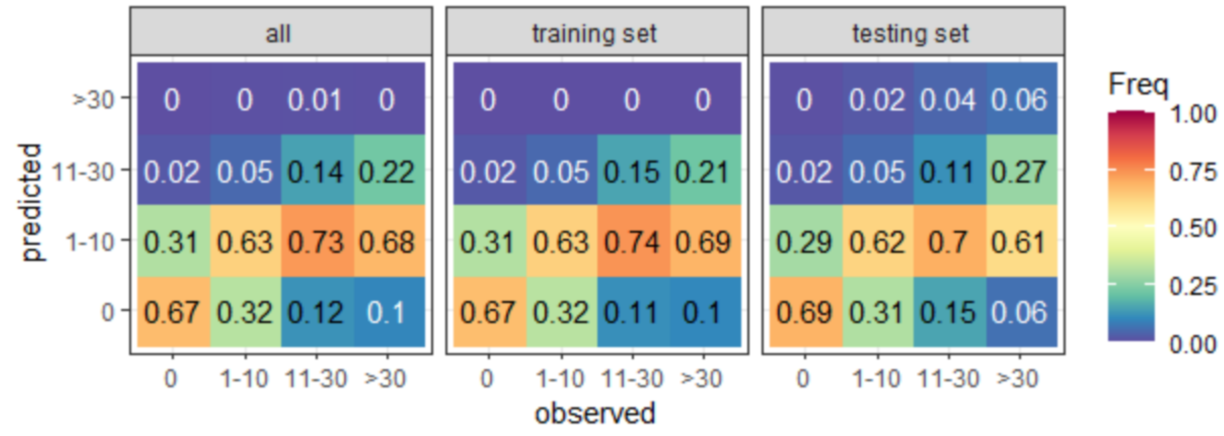
```
temp <- model %>% predict(train_data)



# ROC threshold based on f1

levels <- levels(data$z.cut)

roc <- function(thr, idx){

  preds <- factor(levels[apply(cbind(1, temp) > thr, 1, function(x)
max(which(x)))],

                  levels=levels)

  y_pred=factor(preds[idx], levels = levels)
```

```
  y_true=factor(data$z.cut[idx], levels = levels)

  table <- table(preds=y_pred, obs=y_true)
  recall <- diag(table) / (apply(table, 1, sum) + 1E-9)
  precision <- diag(table) / (apply(table, 2, sum) + 1E-9)
  f1 <- mean(2 * (recall*precision) / (recall+precision + 1E-9))

  rmse <- mean(sapply(1:length(table(as.numeric(y_true))),
                      function(i) sqrt(mean(i -
as.numeric(y_pred)[as.numeric(y_true)==i])^2)))

  c(recall=mean(recall), precision=mean(precision), f1=f1, rmse=rmse)
}
thr <- seq(0.01, .99, by=.01)
res <- data.frame(thr=rep(thr, 3),
                  rbind(do.call("rbind", lapply(thr, roc, 1:nrow(data))),
                        do.call("rbind", lapply(thr, roc, who_train)),
                        do.call("rbind", lapply(thr, roc, who_test))),
                  what=rep(c("total", "trained", "test"), each=length(thr)))
best <- which.max(filter(res, what=="trained")$f1)
best <- which.min(filter(res, what=="trained")$rmse)

temp <- model %>% predict(train_data)

best.thr <- .5
levels <- levels(data$z.cut)

preds <- factor(
  levels[apply(cbind(1, temp) > best.thr, 1, function(x)
max(which(cumprod(x)==1)))],
                levels=levels(data$z.cut))
```

```
table(preds) / table(data$z.cut)


# performance

mean(preds == data$z.cut)

mean(preds[who_test] == data$z.cut[who_test])

mean(abs(as.numeric(data$z.cut)[who_test] - as.numeric(preds)[who_test]) <= 1)

mean(abs(as.numeric(data$z.cut)[who_test] - as.numeric(preds)[who_test]) <= 2)


# Confusion matrix weighted by observations

table <- table(preds=preds, obs=data$z.cut)

conf.mat <- data.frame(t(t(table)/apply(table, 2, sum)), what="all")

table <- table(preds=preds[who_train], obs=data$z.cut[who_train])

conf.mat <- rbind(conf.mat, data.frame(t(t(table)/apply(table, 2, sum)),
what="training set"))

table <- table(preds=preds[who_test], obs=data$z.cut[who_test])

conf.mat <- rbind(conf.mat, data.frame(t(t(table)/apply(table, 2, sum)),
what="testing set"))

colnames(conf.mat) <- c("predicted", "observed", "Freq", "what")
```

Fig. 3.15 Fig. 3.16

```
p <- ggplot(conf.mat) + geom_raster(aes(x=observed, y=predicted, fill=Freq)) +

  scale_fill_gradientn(colours=rev(myPalette(100)), limits=c(0,1)) + theme_bw()
+

  scale_y_discrete(drop=FALSE) +

  facet_wrap(~what) + coord_equal() + geom_text(aes(x=observed, y=predicted,
label=round(Freq,2)),



colour=c("white", "black")[(conf.mat$Freq > .1)+1])

p + scale_x_discrete(labels=c("0", "1-10", "11-30", ">30")) +
scale_y_discrete(labels=c("0", "1-10", "11-30", ">30"))



p
```



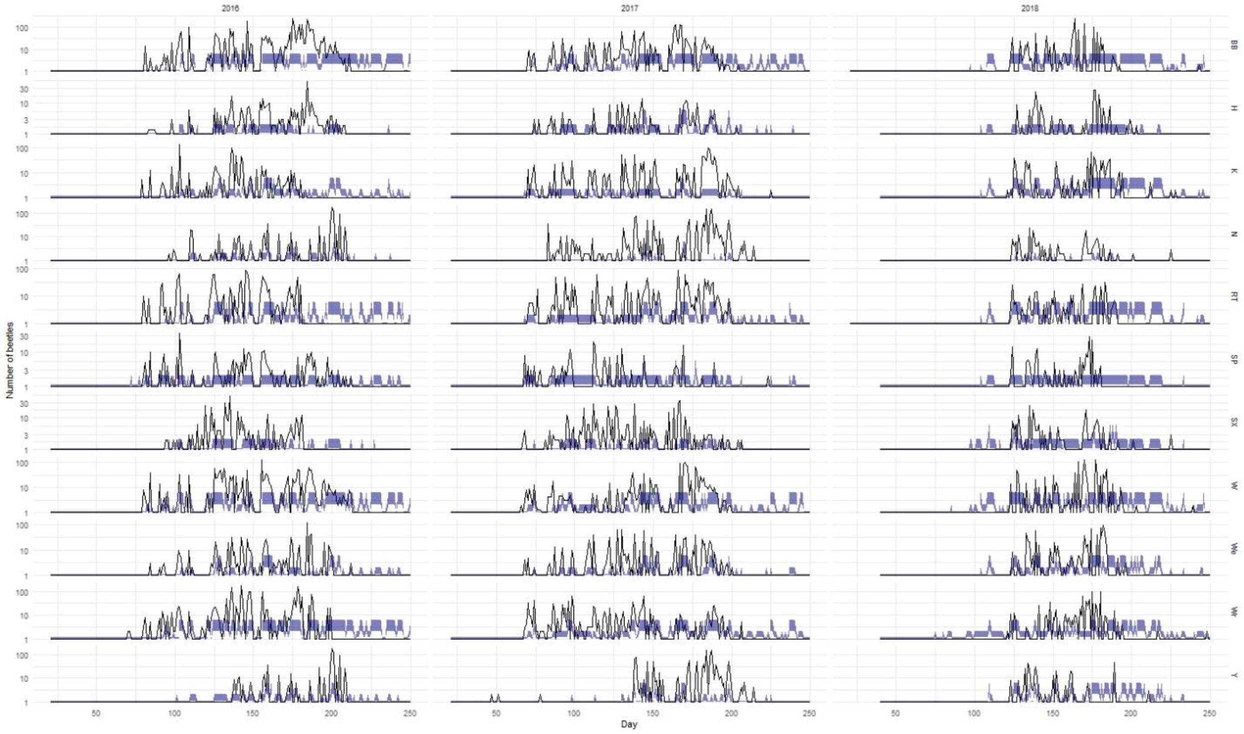
...and test forecasting ability by hindcasting against the test dataset:

```
preds_df <- cbind(data, preds)


preds_df$lo <- c(0,0.1,1.1,5.1,20.1)[as.numeric(preds_df$preds)]

preds_df$hi <- c(.1,1.1,5.1,20.1,max(preds_df$z)+10)[as.numeric(preds_df$preds)]


preds_df$trained <- !(1:nrow(preds_df) %in% who_test)

preds_df <- merge(preds_df, traps)

r <- filter(preds_df, year > 2015) %>%
```

Fig. 3.16

```
  ggplot + geom_line(aes(x=day, y=z+1, group=x), colour="black", show.legend =
FALSE) +

  geom_ribbon(aes(x=day, ymin=lo+1, ymax=hi+1), alpha=.5, fill="navyblue") +
theme_minimal(base_size=8) +

  facet_grid(trap~year, scales = "free") + scale_y_log10()


r+ggtitle("") +
```

```
  xlab("Day") + ylab("Number of beetles")
```



Appendix A, Chapter 4

This Chapter attempted to correlate suction-trap data with field-level data from sticky-traps. It also examined the spatial synchrony of populations recorded from both trap networks.

Libraries

The following libraries are required for this Chapter:

```
library(mgcv)

library(ncf)

library(Hmisc)

library(itsadug)

library(RColorBrewer)
library(tidyr)
```

The tidyr package was used to format the data files when required, the RColorBrewer package was used for graphical output, the itsadug package was used to format table output. The Hmisc package was used for data exploration. The statistical analyses were performed using the mgcv, ncf and itsadug packages

Data

The following data files will be used. These data are available on request.

```
setwd("P:/Final datasets")
DATA<-read.csv("Field weekly data coded means.csv")
CORRDATA <-read.csv("Correlation workbook 2019.csv")
DFS<-read.csv("Field and trap weekly data.csv")
TCI<-read.csv("Trap Count Index.csv")
UKcoast <- read.csv("UK coastline R.csv")
spline <- read.csv("SplineAllYears.csv")
splineSTALL <- read.csv("All Sites Daily All Years v3.csv")
```

Statistical distribution of the data

All data were interrogated with the fitdistrplus package to determine the best statistical distribution. See Appendix A, Chapter 2 for more detail on this step.

Filtering data using the Central England Temperature data in an attempt to overcome the asynchronous sampling issue.

The Central England Temperature (CET) data series was downloaded from https://www.metoffice.gov.uk/hadobs/hadcet/

A converter was used to transform the CET matrix into a csv file. The converter is available here: http://rstudio-pubs-static.s3.amazonaws.com/2117_19f0a65e216747f38f13c5318ce546de.html.

CET converter

```r
#Function to convert data to tabular form
transform.cet <- function(url) {
  # Libraries
  require(reshape2)
  require(lubridate)  # Excellent package for working with dates
  # Read file
  cet <- NULL  # reset object in case already exists
  cet <- read.table(url, quote = "\"")
  # Rename header
  var_names <- c("year", "day", "01", "02", "03", "04", "05", "06", "07",
                 "08", "09", "10", "11", "12")
  colnames(cet) <- var_names
  # Transpose data
  cet.melt <- melt(cet, c("year", "day"))
  cet.melt <- subset(cet.melt, value != -999)  # remove unvalid dates

  cet.melt$date <- dmy(paste(cet.melt$day, "-", cet.melt$variable, "-",
cet.melt$year,
                             sep = ""))
  cet.melt$temperature <- cet.melt$value
  return(cet.melt[, c("date", "temperature")])
}

#Process the files direct from URLs (check URLs remain the same!)
cet.mean <-
transform.cet("http://www.metoffice.gov.uk/hadobs/hadcet/cetdl1772on.dat")

cet.min <-
transform.cet("http://www.metoffice.gov.uk/hadobs/hadcet/cetmindly1878on_urbadj4
.dat")
cet.max <-
transform.cet("http://www.metoffice.gov.uk/hadobs/hadcet/cetmaxdly1878on_urbadj4
.dat")

#Rename
names(cet.mean)[names(cet.mean) == "temperature"] <- "mean_temp"
names(cet.min)[names(cet.min) == "temperature"] <- "min_temp"
names(cet.max)[names(cet.max) == "temperature"] <- "max_temp"

#export to csv
write.csv(cet.mean, "cet_mean.csv", row.names = F)
write.csv(cet.min, "cet_min.csv", row.names = F)
write.csv(cet.max, "cet_max.csv", row.names = F)

#NOTE, Temperature is expressed in tenths of a degree C so need to divide by 10
if you require centigrades
```

Once converted, the data were used to filter dates from the original sticky- and suction-trap datasets. All dates below 11 degrees Centigrade were removed and the count samples were divided across the remaining days to produce a weekly mean count for all sites in all years.

This was done manually in Excel.

The resulting file was then used to compare each standard week's suction-trap data with sticky-trap data from the same week. A simple Pearson's correlation was run for each site/year combination, for example:

```
cor.test(DATA$Rothamsted2008a,DATA$S2008_02, na.action = "na.exclude")

##
##  Pearson's product-moment correlation
##
## data:  DATA$Rothamsted2008a and DATA$S2008_02
## t = -0.056188, df = 4, p-value = 0.9579
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  -0.8209343  0.8017511
## sample estimates:
##        cor
## -0.0280829
```

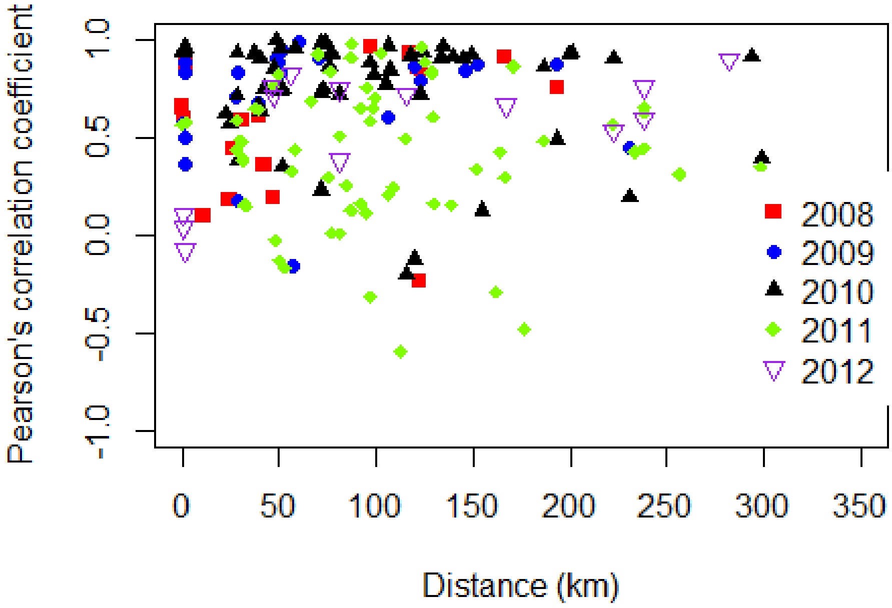
Where, "Rothamsted2008a" is the weekly means for the Rothamsted suction-trap in 2008, and "s2008_X" is the weekly means for 2008 at field site no. X.

The output from these correlations were assembled into a spread sheet and then plotted.

Fig. 4.3

```
#Convert Year to factor
attach(CORRDATA)
CORRDATA$Year <- factor(CORRDATA$Year, ordered = TRUE)
#split into constituent years aslist "X"
X<-split(CORRDATA, Year, drop=TRUE)
#list those years in Global Environment
names(X)<-c("2008", "2009", "2010", "2011", "2012")
list2env(X, envir=.GlobalEnv)

## <environment: R_GlobalEnv>

detach(CORRDATA)
#plot all four years on same graph
attach(`2008`)
plot(Distance, Correlation.coefficient, col="red", pch=15, xlab="Distance (km)",
ylab= "Pearson's correlation coefficient", xlim=c(0,350), ylim=(c(-1, 1)))
detach (`2008`)
attach(`2009`)
points(Distance, Correlation.coefficient, col="blue", pch=16)
detach (`2009`)
attach(`2010`)
points(Distance, Correlation.coefficient, col="black", pch=17)
detach(`2010`)
attach(`2011`)
```

```
points(Distance, Correlation.coefficient, col="chartreuse", pch=18,
xlim=c(0,350))
detach(`2011`)
attach(`2012`)
points(Distance, Correlation.coefficient, col="purple", pch=25)
detach (`2012`)
legend(x=290, y=0.32, legend=c("2008", "2009", "2010", "2011", "2012"),
col=c("red", "blue", "black", "chartreuse", "purple"), pch = c(15, 16, 17, 18,
25), box.lty=0)
```



Visualising and modelling spatial autocorrelation

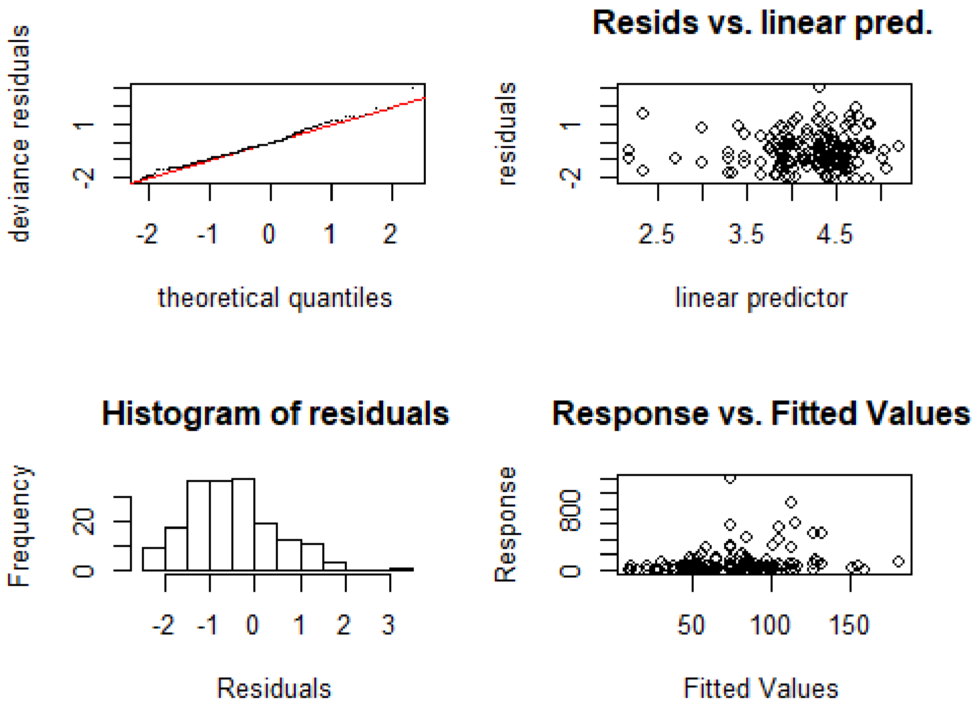
Section 4.3.2.1

Once the data had been assembled and their statistical distribution assessed a GAM was run on each week's data across all field traps to determine the spatial structure of the data.

For brevity only two weeks are shown here.

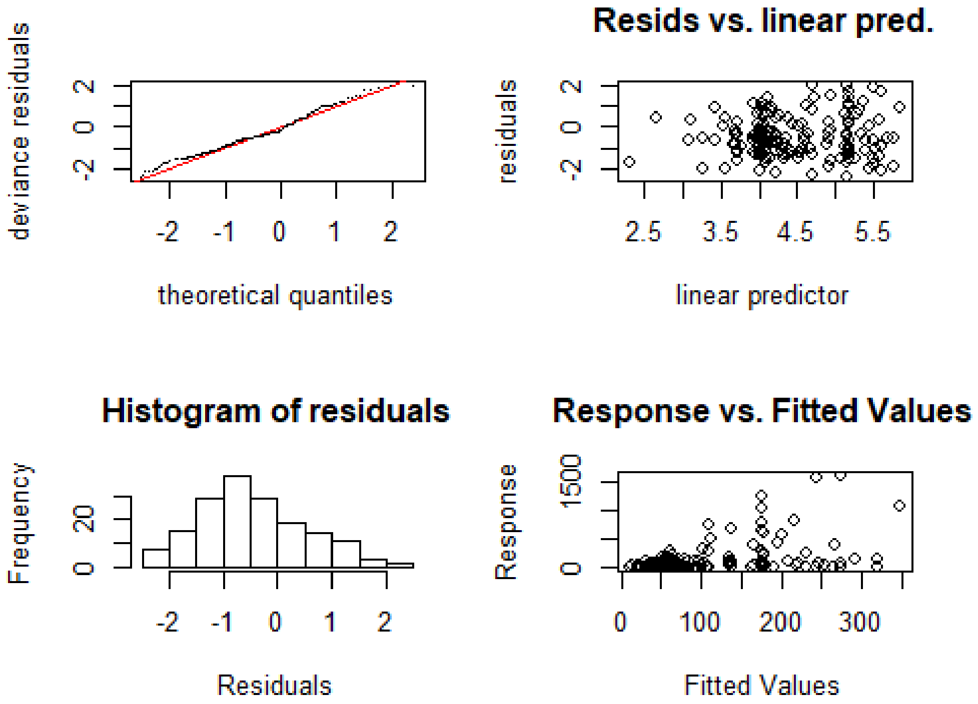
```
myPalette <- colorRampPalette(brewer.pal(11, "Spectral"))
# Run a model with Lat and Long inside an s() function. Set the degrees of
freedom with k =.

par(mar=c(5, 4, 4, 2))
par(mfrow=c(2,2))

Week7.gam<- gam(Week7~s(Longitude, Latitude, bs="ts"), data = DFS, family =
```

```
nb())
gam.check(Week7.gam)
```



```
##
## Method: REML   Optimizer: outer newton
## full convergence after 5 iterations.
## Gradient range [0.0001511175,0.0003108326]
## (score 881.2502 & scale 1).
## Hessian positive definite, eigenvalue range [0.8098846,109.7071].
## Model rank =  30 / 30
##
## Basis dimension (k) checking results. Low p-value (k-index<1) may
## indicate that k is too low, especially if edf is close to k'.
##
##                           k'   edf k-index p-value
## s(Longitude,Latitude) 29.00  1.73    0.74    0.26
```

```
Week8.gam<- gam(Week8~s(Longitude, Latitude, bs="ts"), data = DFS, family =
nb())
gam.check(Week8.gam)
```

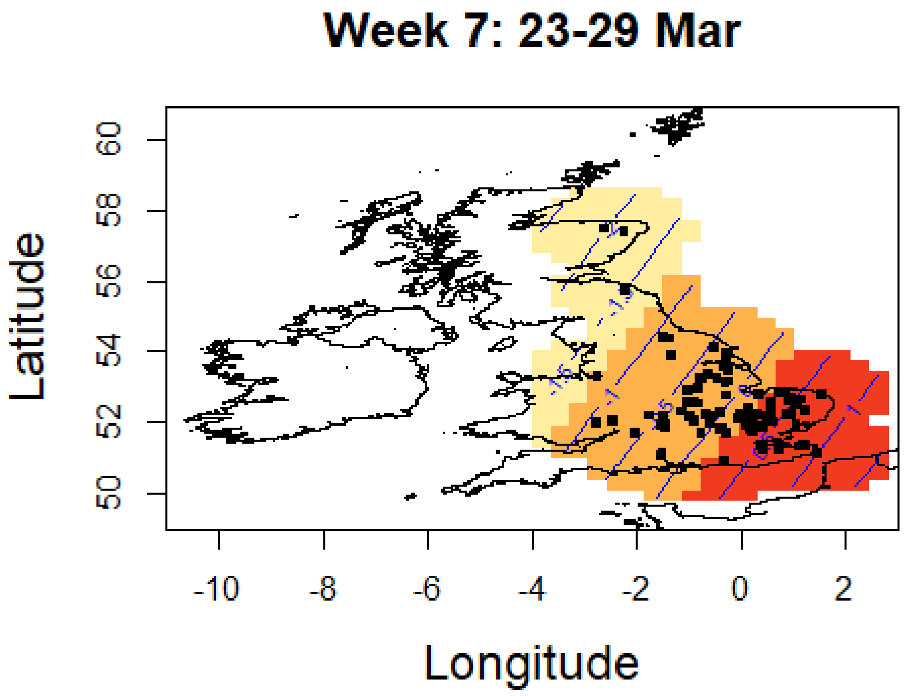


```
##
## Method: REML   Optimizer: outer newton
## full convergence after 4 iterations.
## Gradient range [5.462679e-10,6.122274e-09]
## (score 922.0661 & scale 1).
## Hessian positive definite, eigenvalue range [3.113599,100.6324].
## Model rank =  30 / 30
##
## Basis dimension (k) checking results. Low p-value (k-index<1) may
## indicate that k is too low, especially if edf is close to k'.
##
##                          k'   edf k-index p-value
## s(Longitude,Latitude) 29.0 12.7    0.84     0.6
```

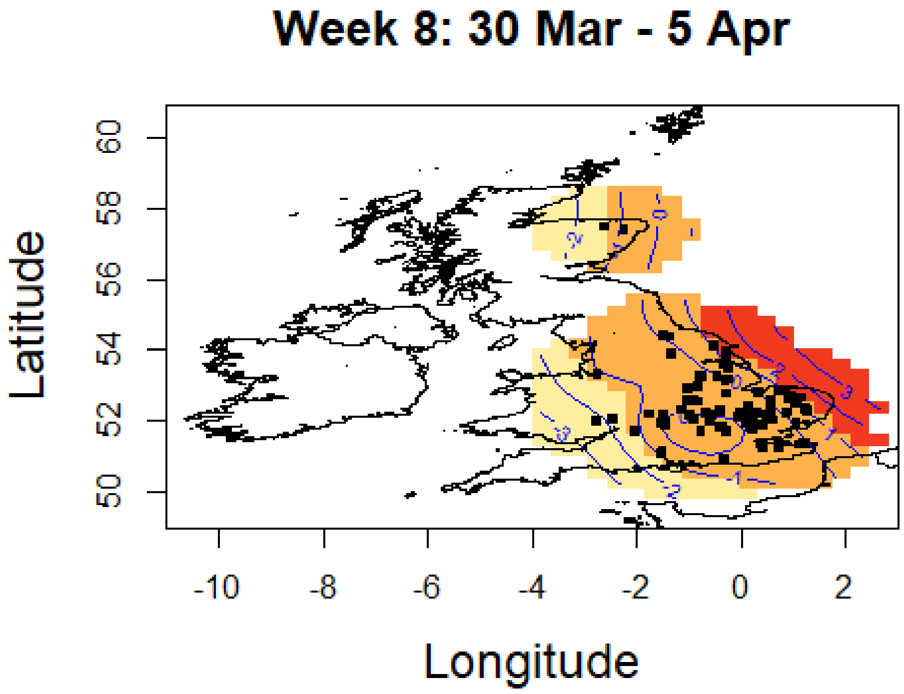
These GAMs were then mapped:

Fig. 4.4

```
par(mfrow=c(1,1))
plot(Week7.gam, select=1, main="Week 7: 23-29 Mar", xlab="Longitude",
     ylab="Latitude", xlim=c(-11,3), ylim=c(49,60.9), cex.main=1.5, cex.lab=1.5,
     cex.axis=1, scheme=2, hcolors=brewer.pal(n = 3, name = "YlOrRd"),
too.far=0.1, cex = 5)
lines(UKcoast)
```



```r
plot(Week8.gam, select=1, main="Week 8: 30 Mar - 5 Apr", xlab="Longitude",
     ylab="Latitude", xlim=c(-11,3), ylim=c(49,60.9), cex.main=1.5, cex.lab=1.5,
     cex.axis=1, scheme=2, hcolors=brewer.pal(n = 3, name = "YlOrRd"),
too.far=0.1, cex = 5)
lines(UKcoast)
```



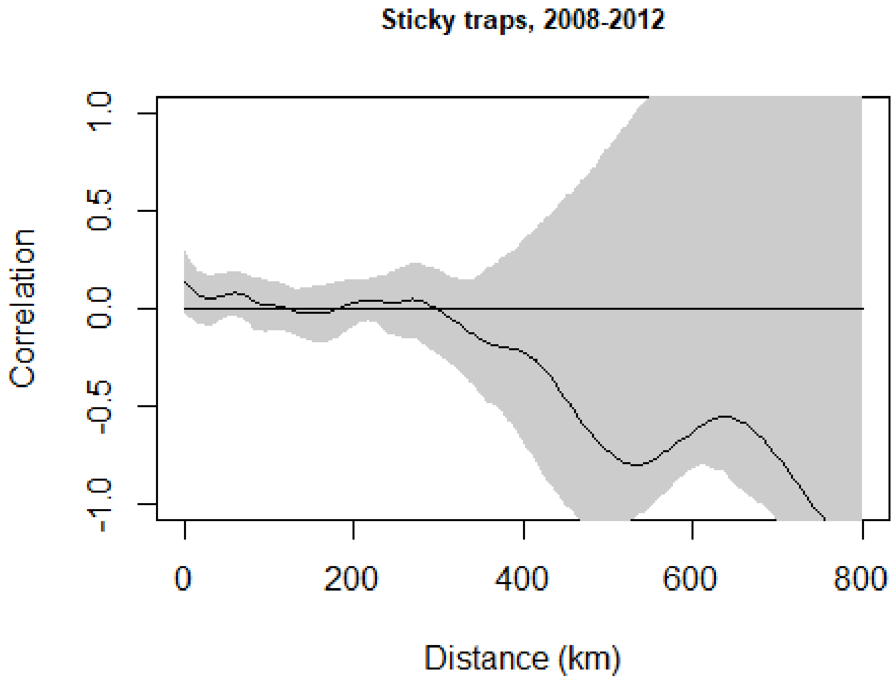
Multivariate spline analysis of field data

Section 4.4.3

The ncf package was used to create spline correlograms of the sticky trap and (below in Section 4.4.4) suction-trap data.

Fig. 4.5

```
attach (spline)
ncols <- dim(spline)[2]
week<-spline[ , 4:ncols]

attach(week)
spline1 <- spline.correlog(x=Longitude, y=Latitude, z=week, resamp = 1100,
                                     latlon = TRUE, xmax=800, na.rm = TRUE)

## Warning in spline.correlog(x = Longitude, y = Latitude, z = week, resamp =
## 1100, : Missing values exist; Pairwise deletion will be used

## Warning in cor(x = x, y = y, use = "pairwise.complete.obs", method =
"pearson"):
## the standard deviation is zero

## 100  of  1100 200  of  1100 300  of  1100 400  of  1100 500  of  1100 600  of
1100 700  of  1100 800  of  1100 900  of  1100 1000  of  1100 1100  of  1100

#summary.spline.correlog(spline)
plot (spline1, xlab="Distance (km)", ylab="Correlation")
title(main="Sticky traps, 2008-2012", cex.main=0.8)
```



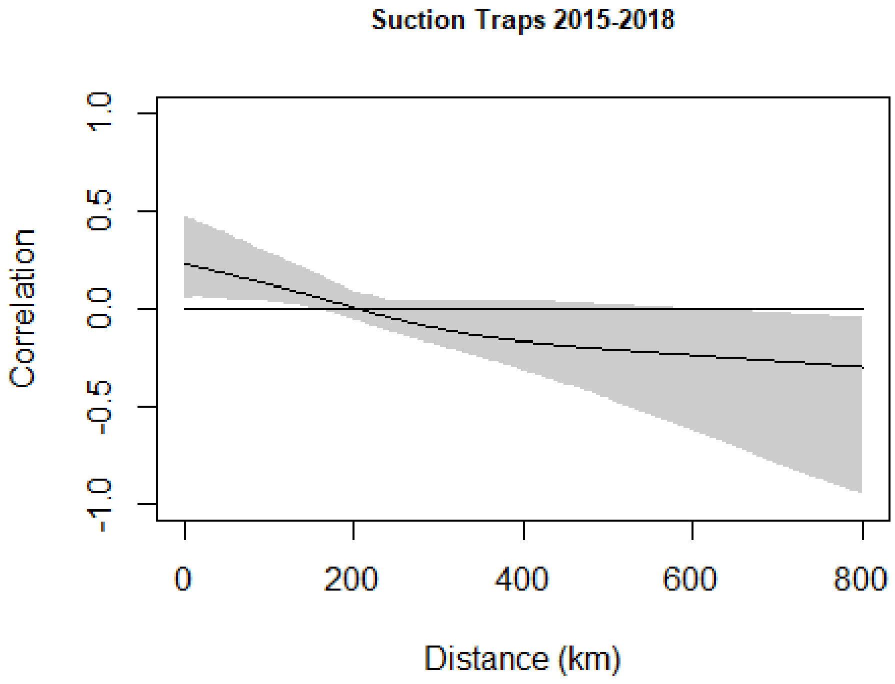
```
summary(spline1)

## $call
## [1] "spline.correlog(x = Longitude, y = Latitude, z = week, resamp = 1100, "
## [2] "    xmax = 800, latlon = TRUE, na.rm = TRUE)"
##
## $estimate
##                 x  e         y
## estimate 121.7491 NA 0.1368644
##
## $quantiles
##                 x         e           y
## 0       0.2742335 0.1365596 -0.13498039
## 0.025   3.6345163 0.1466514 -0.03352307
## 0.25   47.6816768 0.2374781  0.08298438
## 0.5   106.8945790 0.3383967  0.14252336
## 0.75  144.9357542 0.9151758  0.20134528
## 0.975 356.5950913 1.4342770  0.29768896
## 1     761.5778412 1.4919550  0.37997711

detach(spline)
detach(week)
```

Multivariate spline analysis of suction-trap data

Section 4.4.4

Fig. 4.6

```
attach (splineSTALL)
ncols <- dim(splineSTALL)[2]
day<-splineSTALL[ , 5:ncols]
attach(day)
spline2 <- spline.correlog(x=Long, y=Lat, z=day, resamp = 1100,latlon = TRUE,
xmax=800, na.rm = TRUE)

## Warning in spline.correlog(x = Long, y = Lat, z = day, resamp = 1100, latlon
=
## TRUE, : Missing values exist; Pairwise deletion will be used

## 100  of  1100 200  of  1100 300  of  1100 400  of  1100 500  of  1100 600  of
1100 700  of  1100 800  of  1100 900  of  1100 1000  of  1100 1100  of  1100

plot (spline2, xlab="Distance (km)", ylab="Correlation")
title(main="Suction Traps 2015-2018", cex.main=0.8)
```

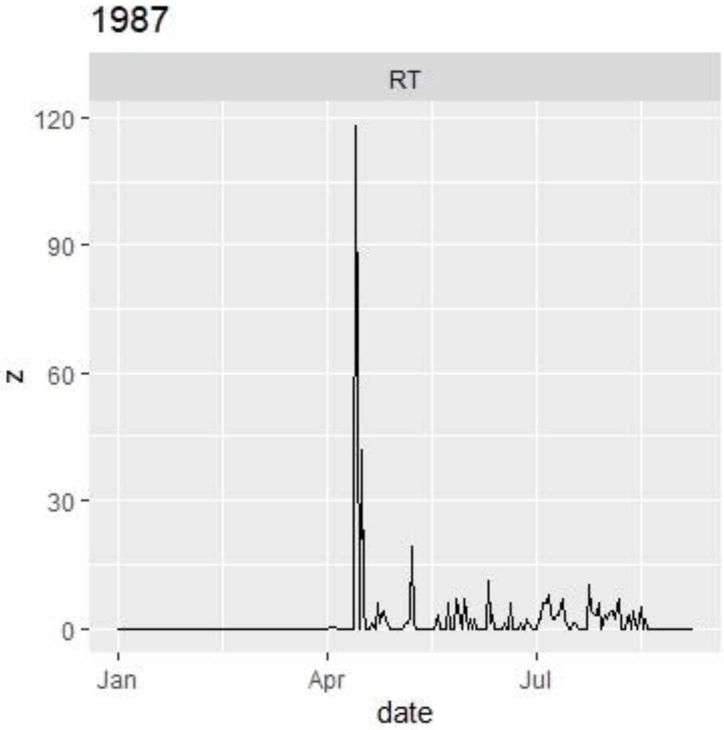


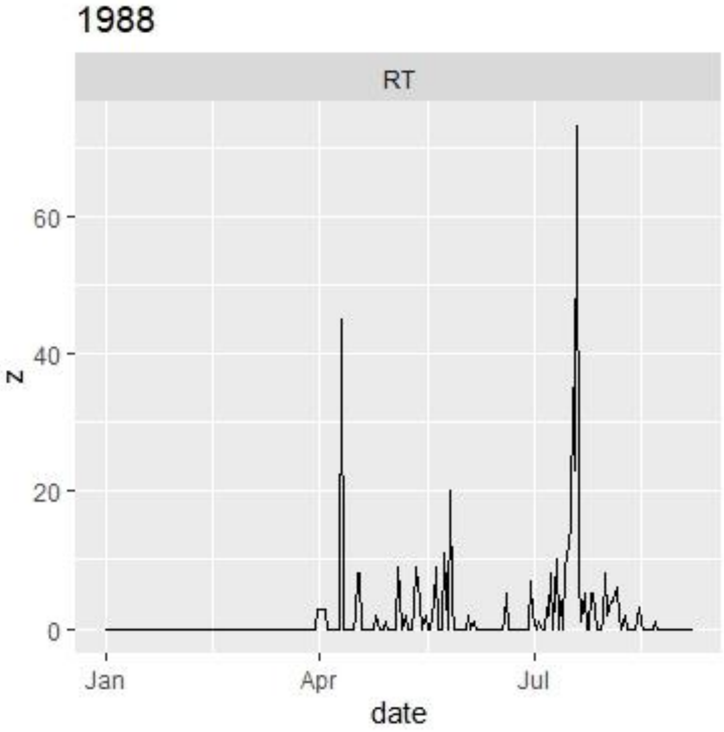
```
summary(spline2)

## $call
## [1] "spline.correlog(x = Long, y = Lat, z = day, resamp = 1100, xmax = 800, "
## [2] "    latlon = TRUE, na.rm = TRUE)"
##
## $estimate
##                 x  e         y
## estimate 204.6627 NA 0.2314857
##
## $quantiles
##              x          e           y
## 0      24.04849   1.258517 -0.08438767
```

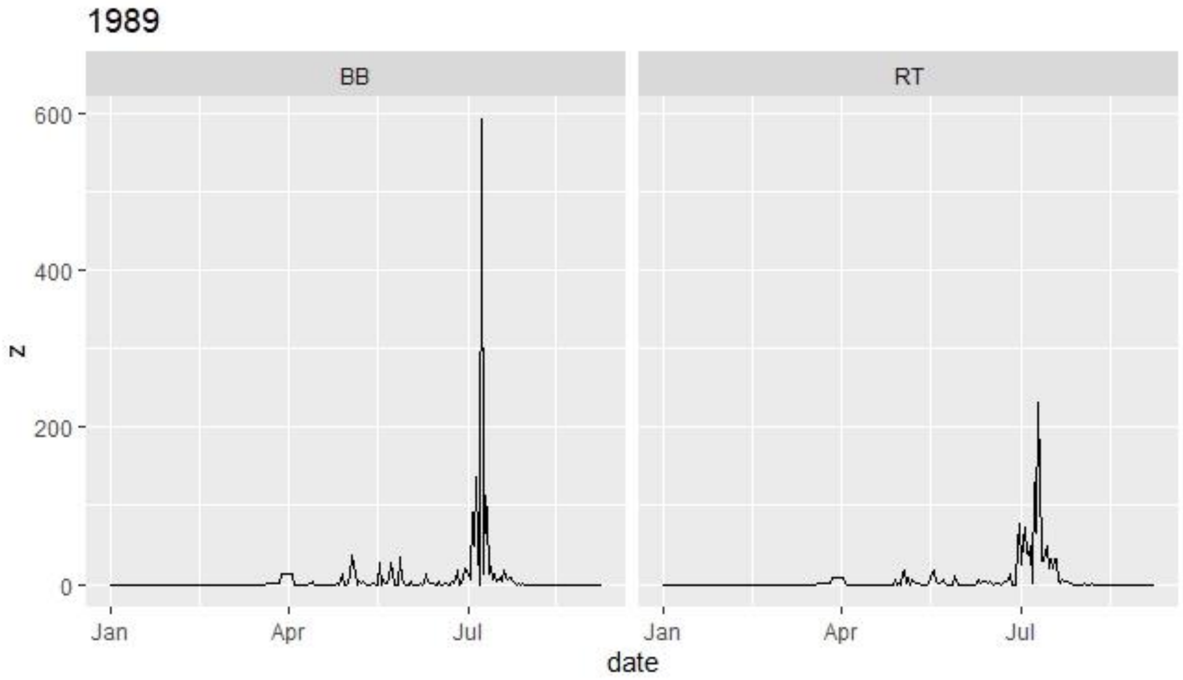
```
## 0.025 155.33379   2.571384  0.05561919
## 0.25  189.15450  13.619554  0.17089633
## 0.5   205.08136  25.813974  0.23244543
## 0.75  223.55064  49.040372  0.30536255
## 0.975 267.36404 126.804332  0.47395594
## 1     666.24178 143.187711  0.99580953
```

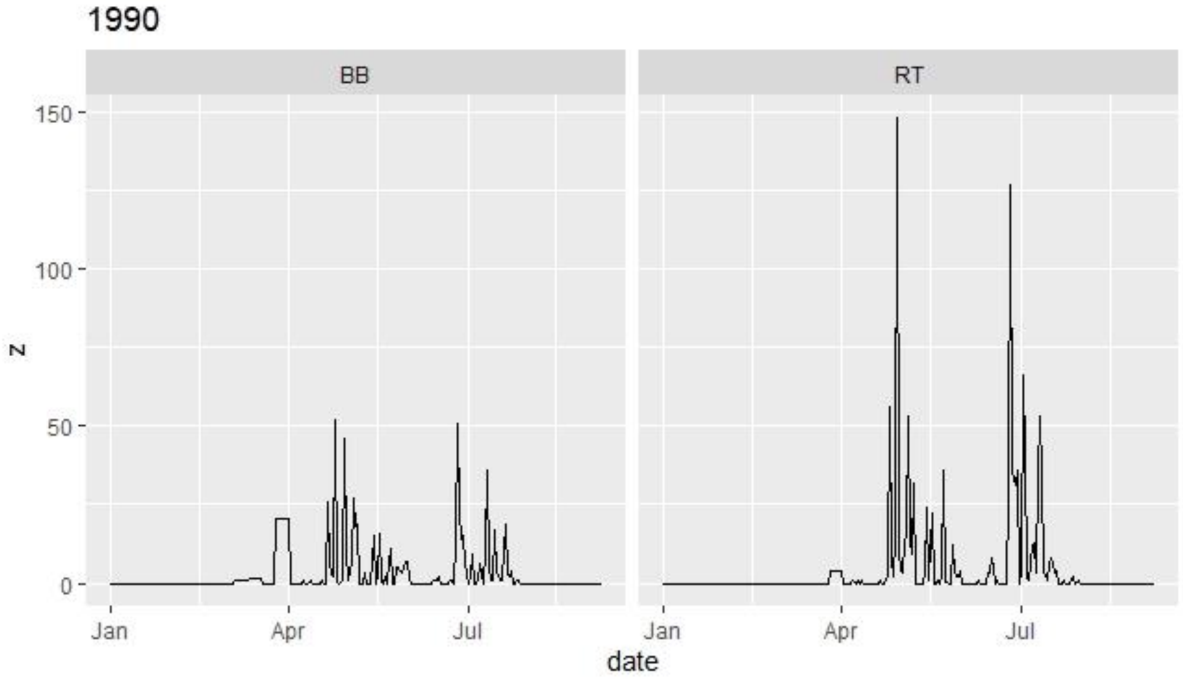
Appendix B: **Daily data plots for all site/years**

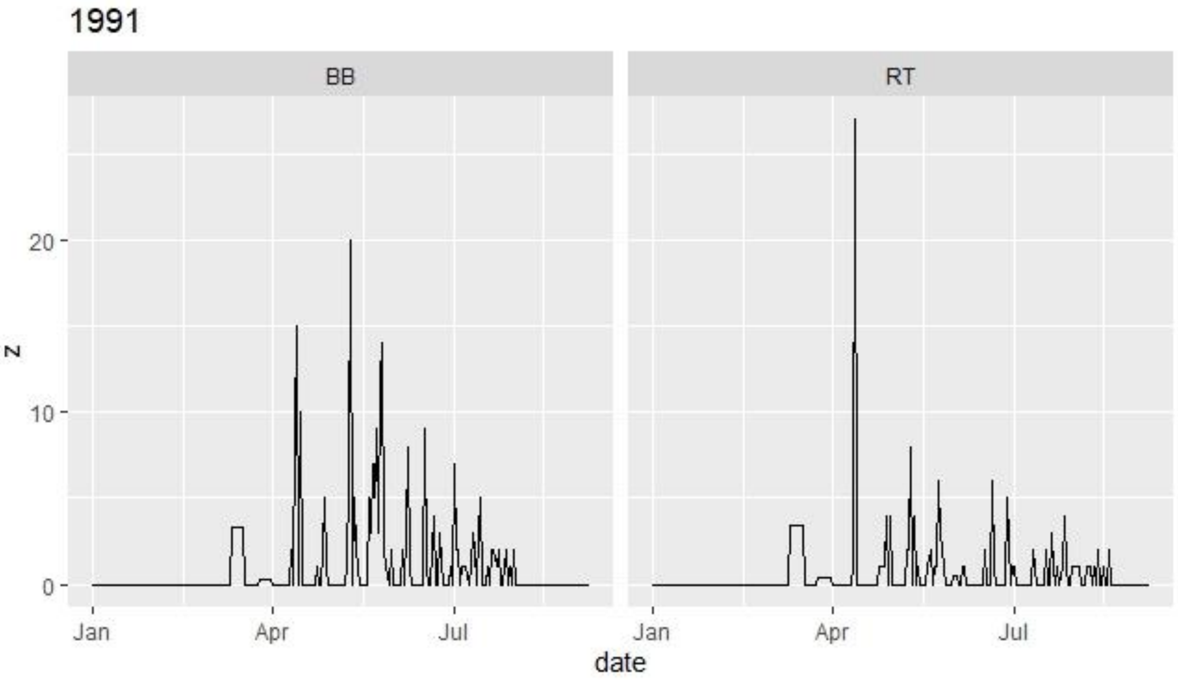
Site Codes: RT: Rothamsted; BB: Broom's Barn; H: Hereford; K: Kirton II; N: Newcastle; SP: Silwood Park; SX: Starcross; W: Wye; We: Wellesbourne; Wr: Writtle; Y: York. In all plots the y-axis (z) is the daily count of *Brassicogethes aeneus.*

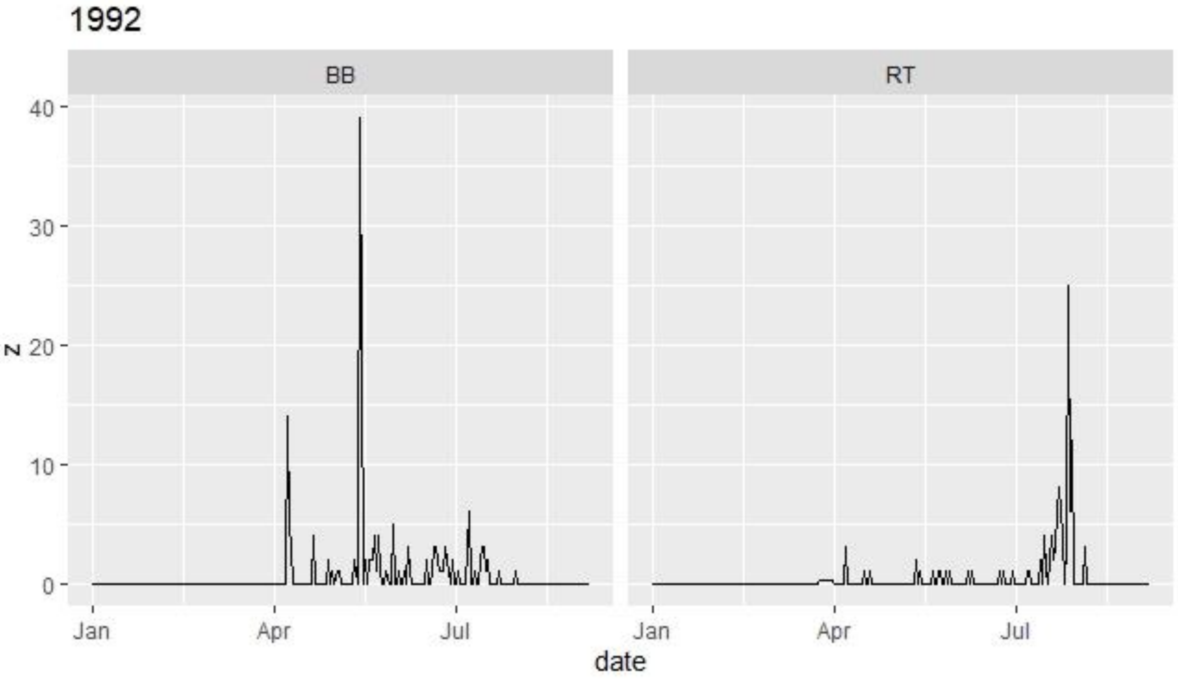


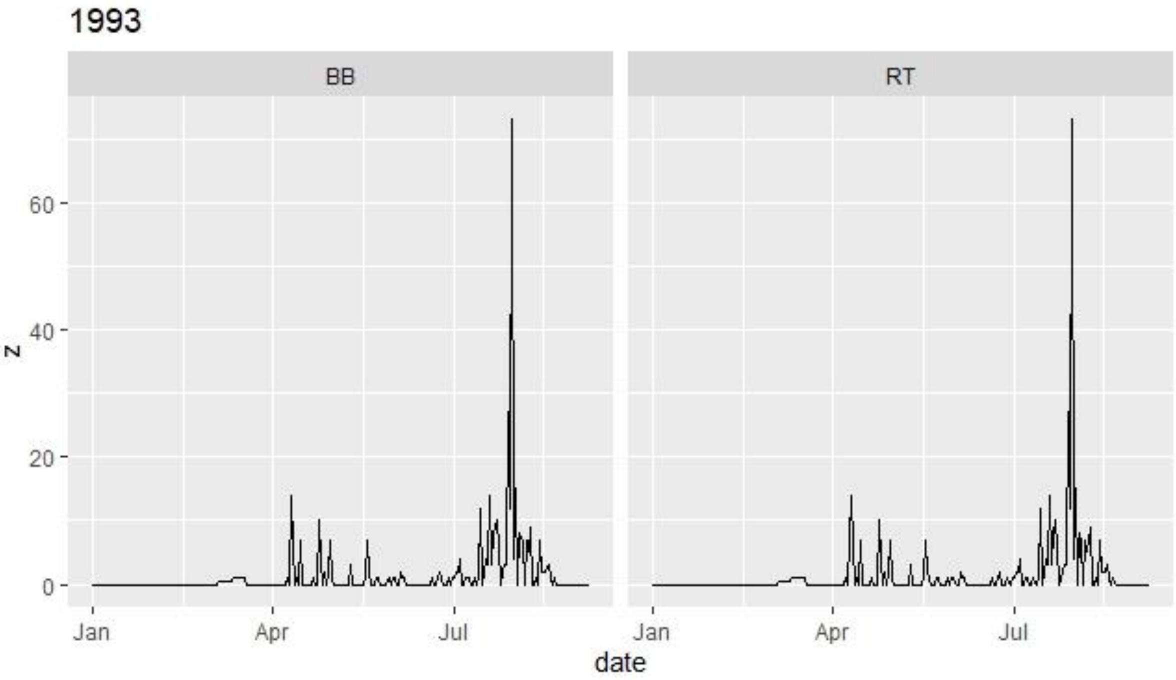


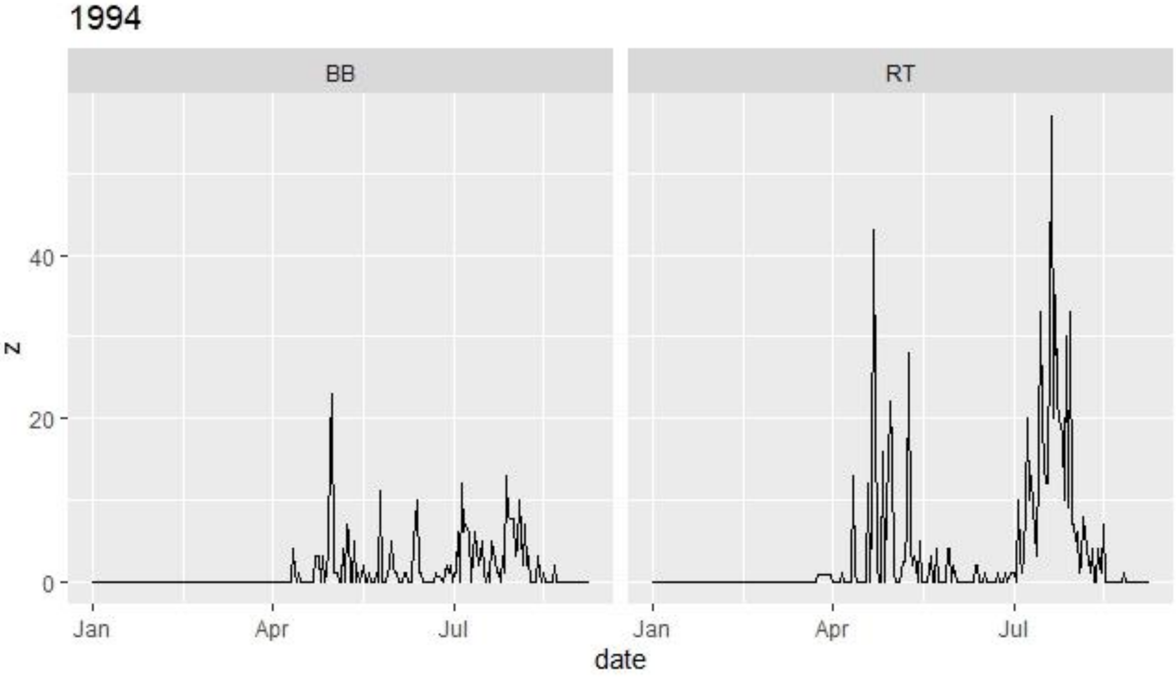

1989
BB
RT
600
400
200
0
z
Jan
Apr
Jul
Jan
Apr
Jul
date

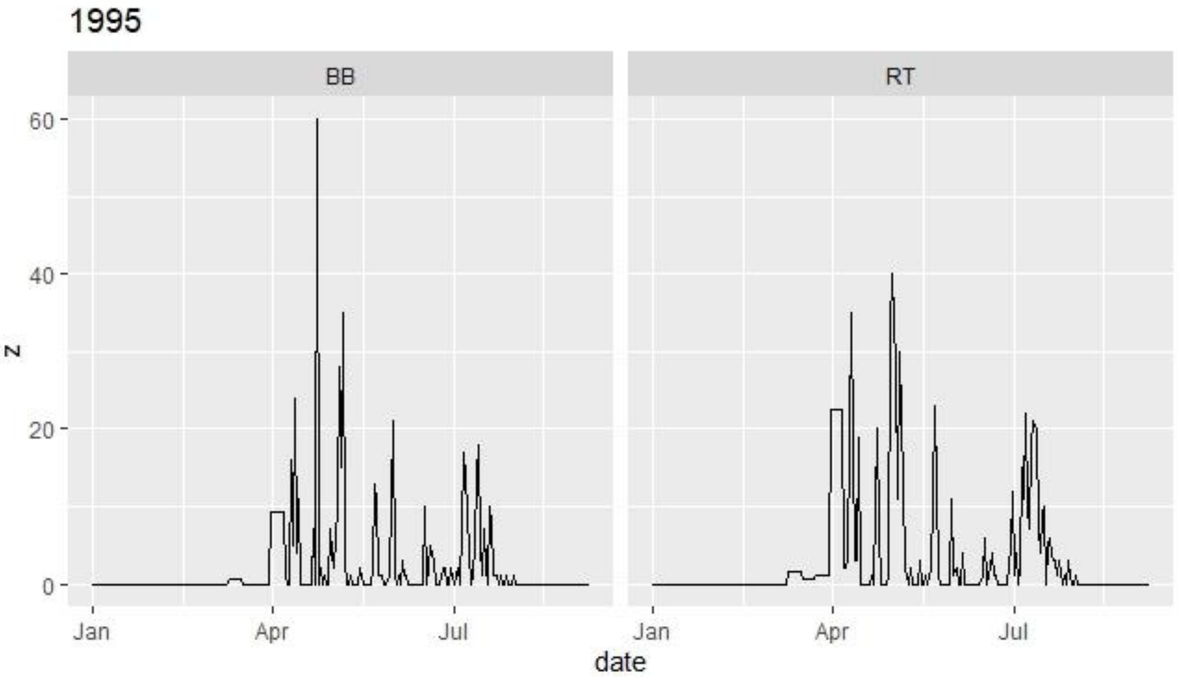

1990
BB
RT
150
100
50
0
z
Jan
Apr
Jul
Jan
Apr
Jul
date

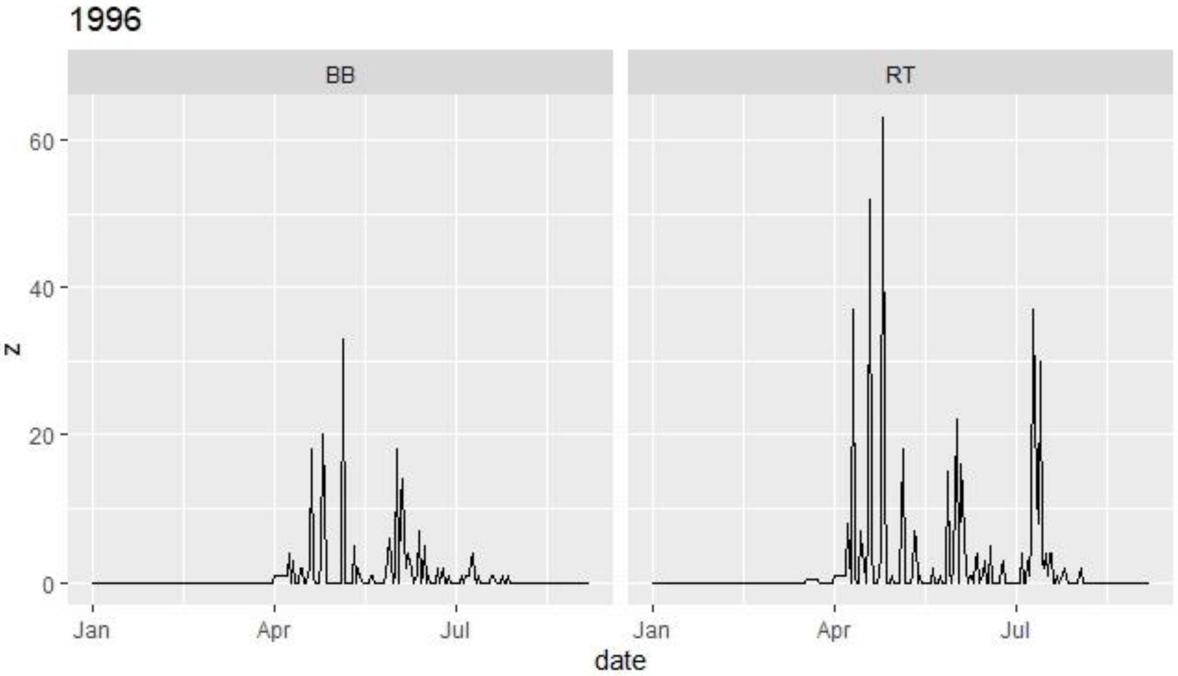

1991
BB
RT
z
0
10
20
Jan
Apr
Jul
Jan
Apr
Jul
date

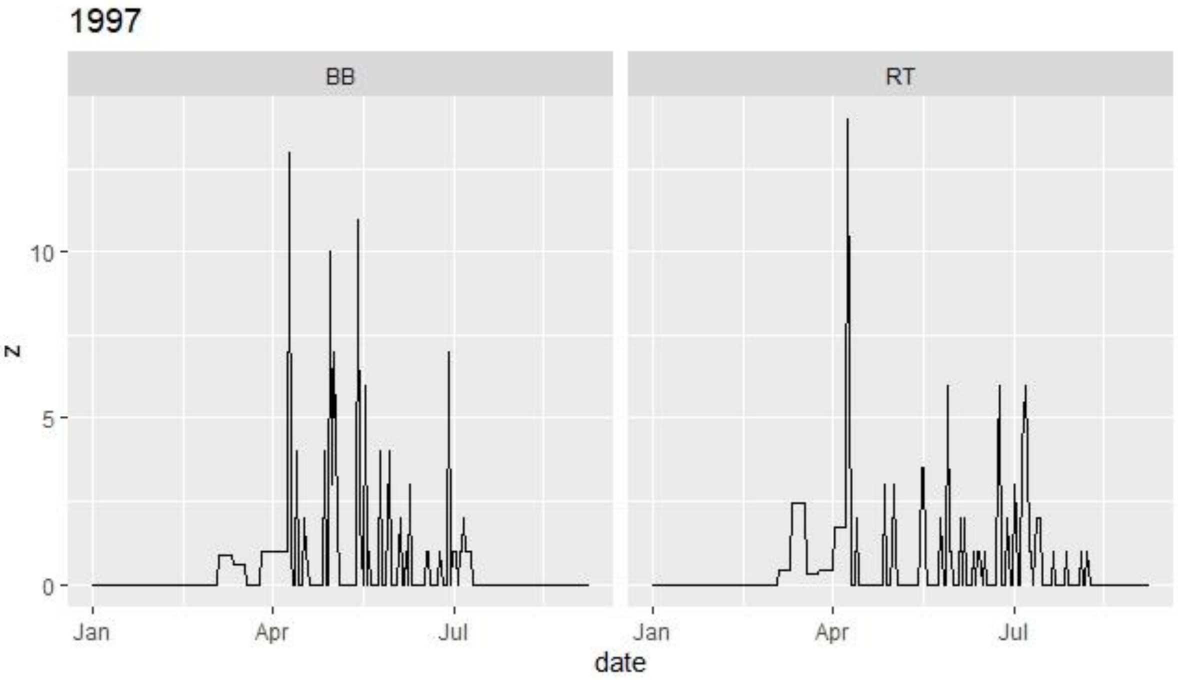

1992
BB
RT
z
0
10
20
30
40
Jan
Apr
Jul
Jan
Apr
Jul
date

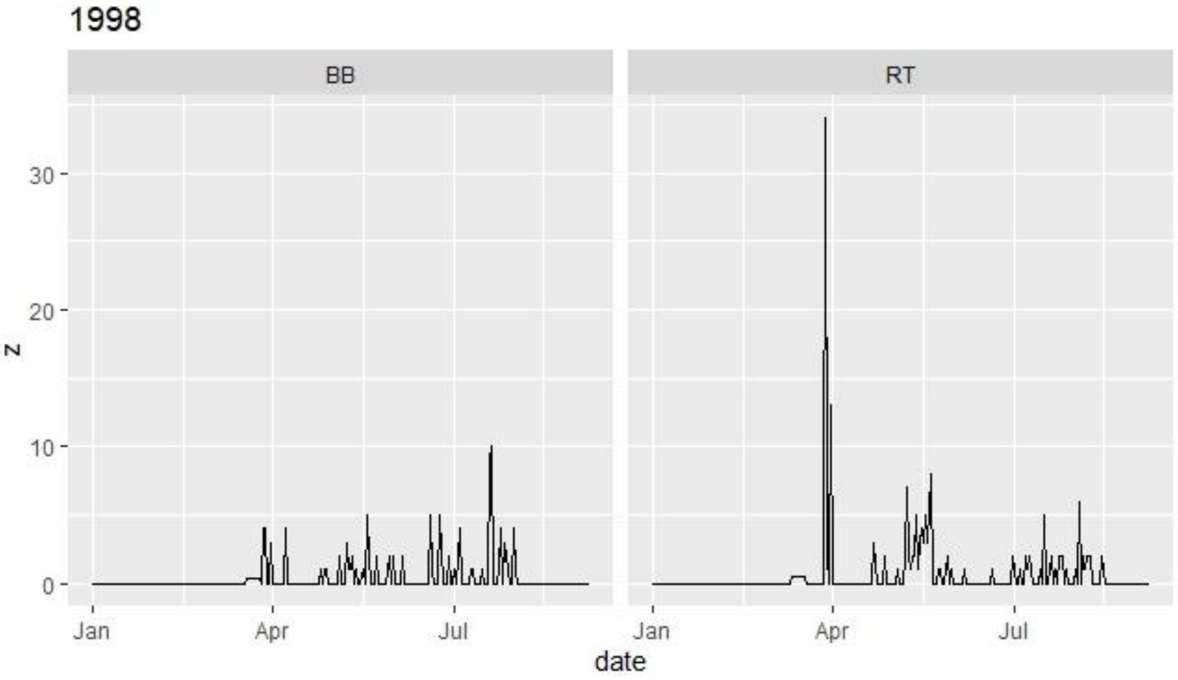

1993
BB
RT
z
60
40
20
0
Jan
Apr
Jul
Jan
Apr
Jul
date

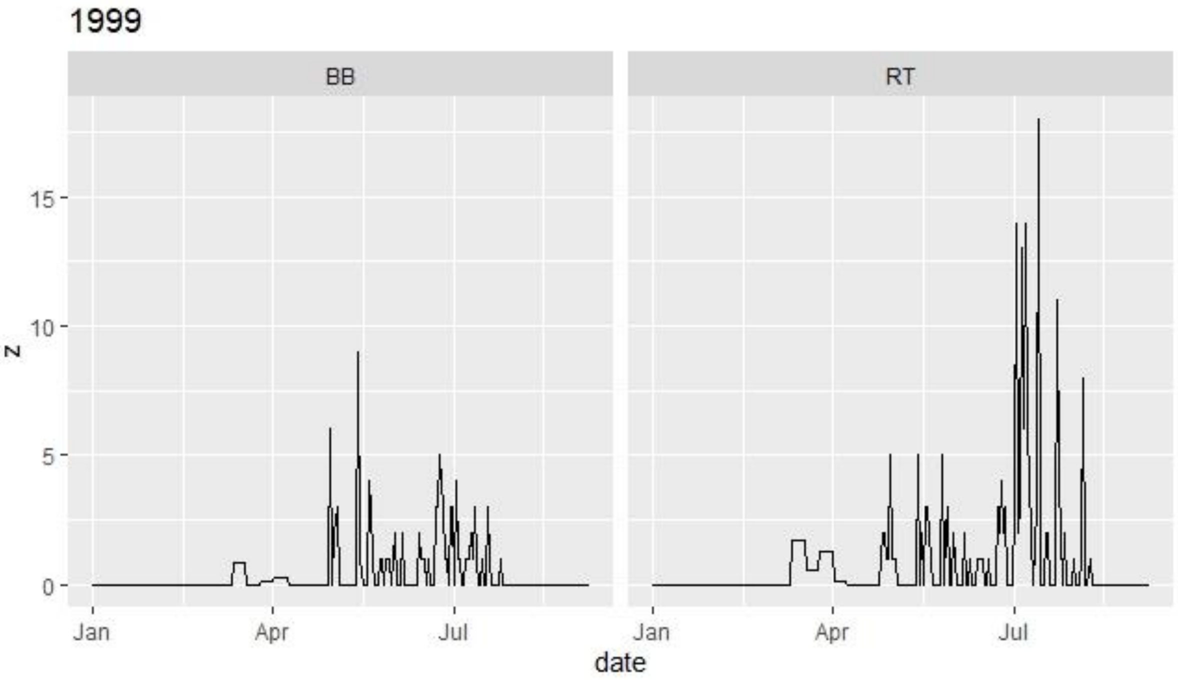

1994
BB
RT
z
40
20
0
Jan
Apr
Jul
Jan
Apr
Jul
date

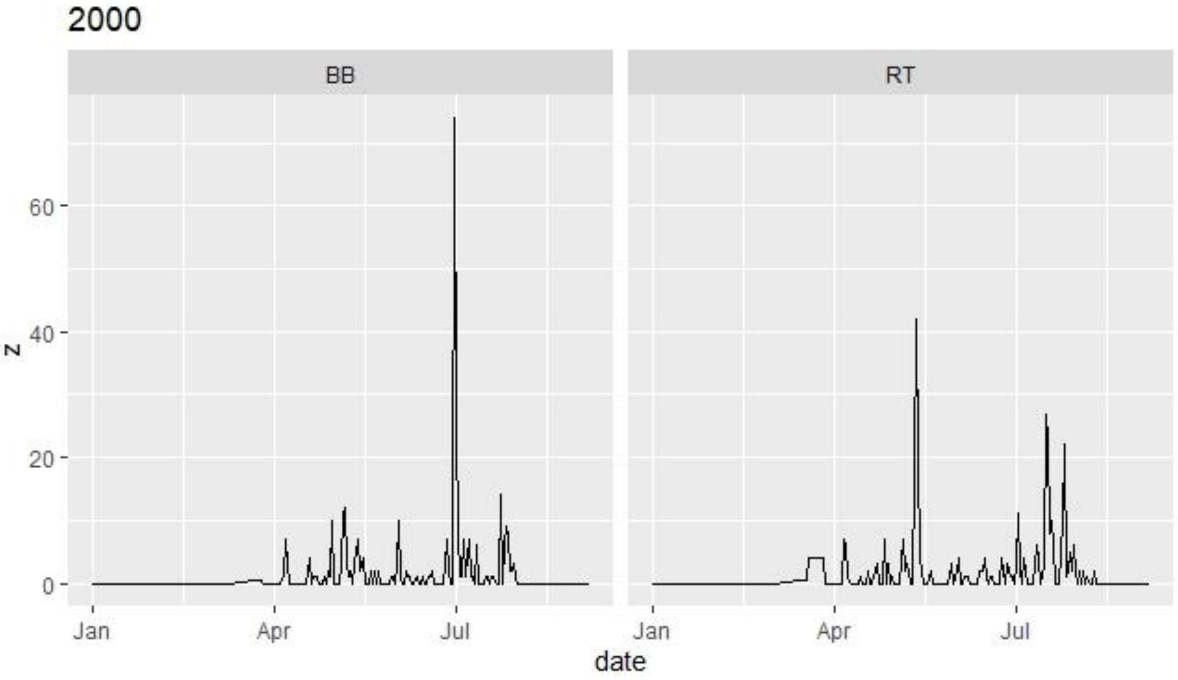

1995
BB
RT
z
0
20
40
60
Jan
Apr
Jul
Jan
Apr
Jul
date

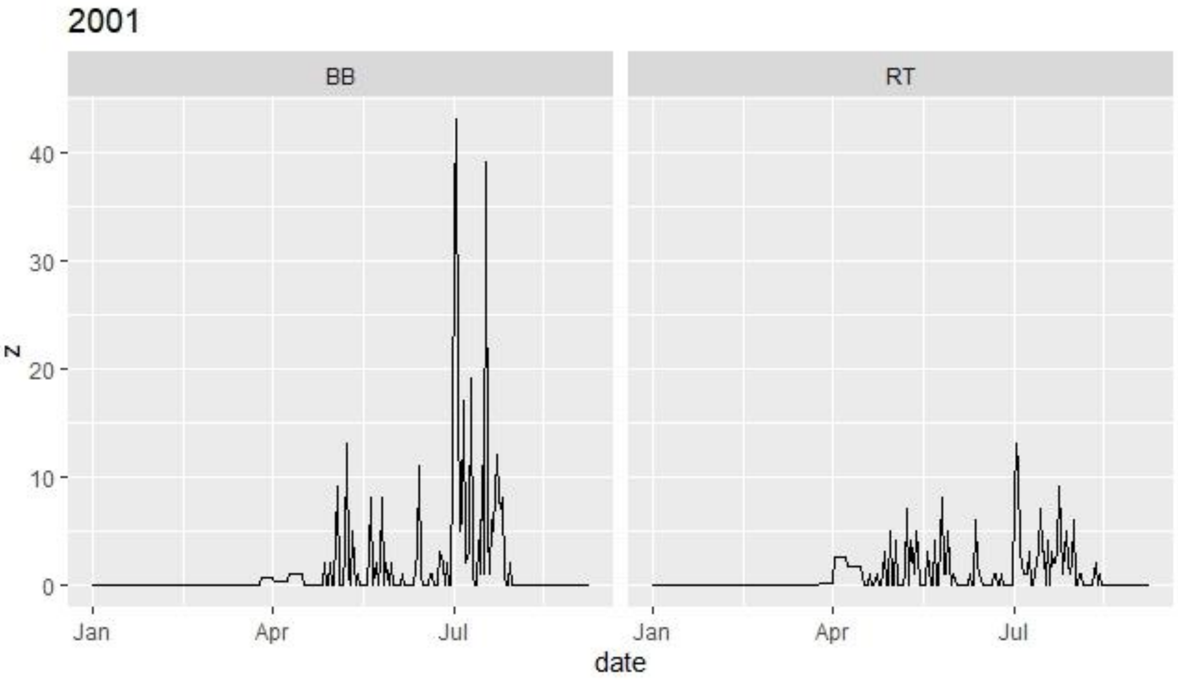

1996
BB
RT
z
0
20
40
60
Jan
Apr
Jul
Jan
Apr
Jul
date

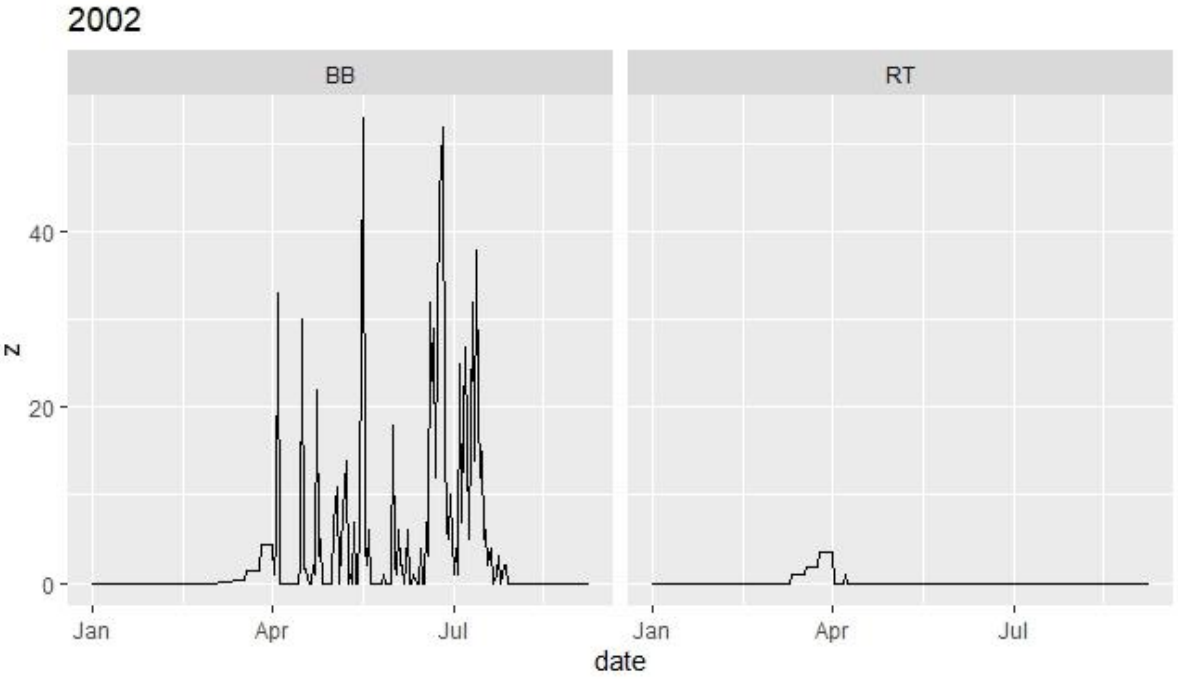

1997
BB
RT
z
0
5
10
Jan
Apr
Jul
Jan
Apr
Jul
date

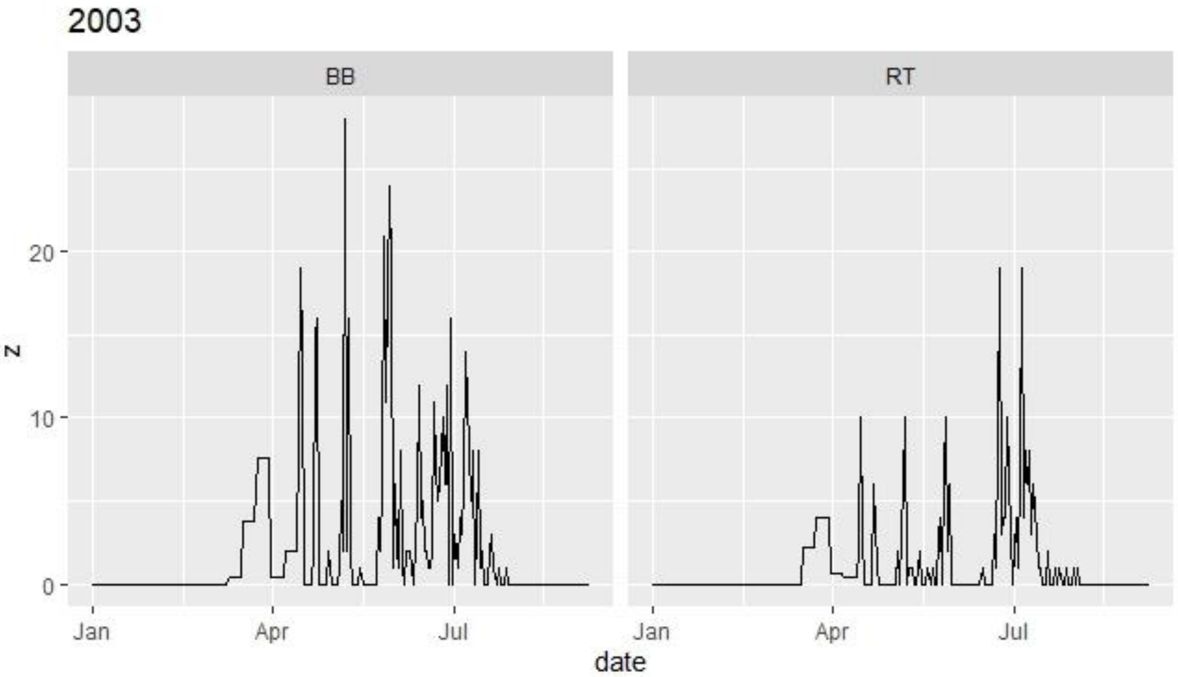

1998
BB
RT
z
0
10
20
30
Jan
Apr
Jul
Jan
Apr
Jul
date

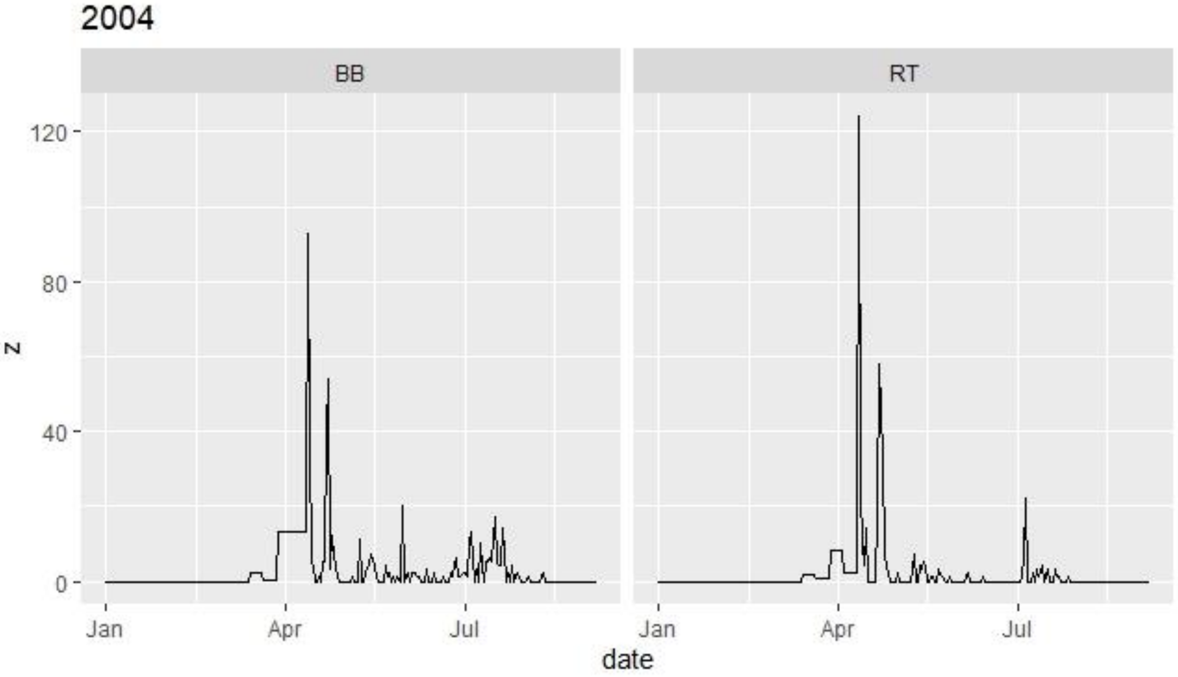

1999
BB
RT
z
0
5
10
15
Jan
Apr
Jul
date

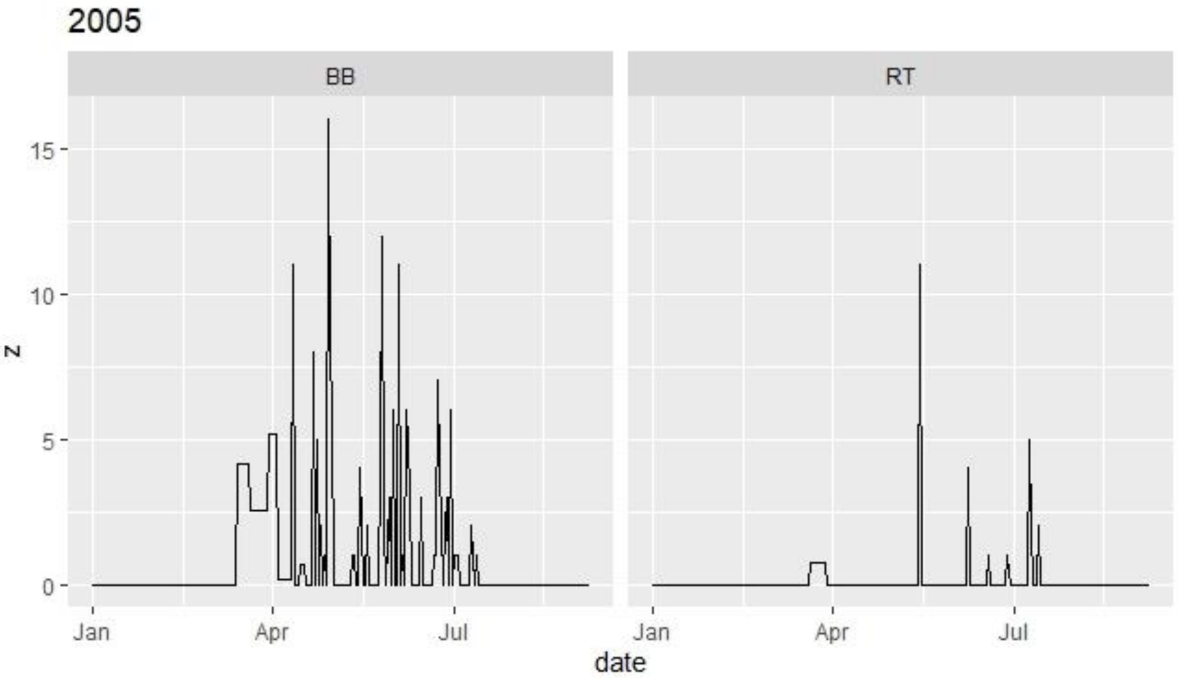

2000
BB
RT
z
0
20
40
60
Jan
Apr
Jul
date

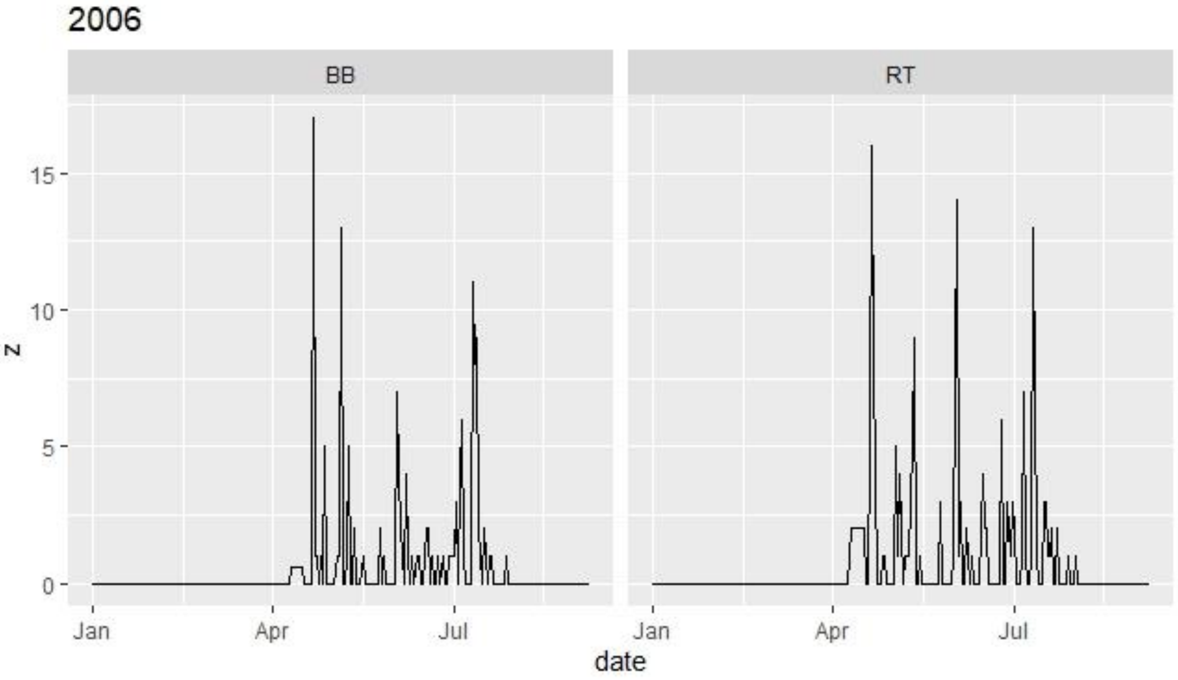

2001
BB
RT
z
40
30
20
10
0
Jan
Apr
Jul
Jan
Apr
Jul
date

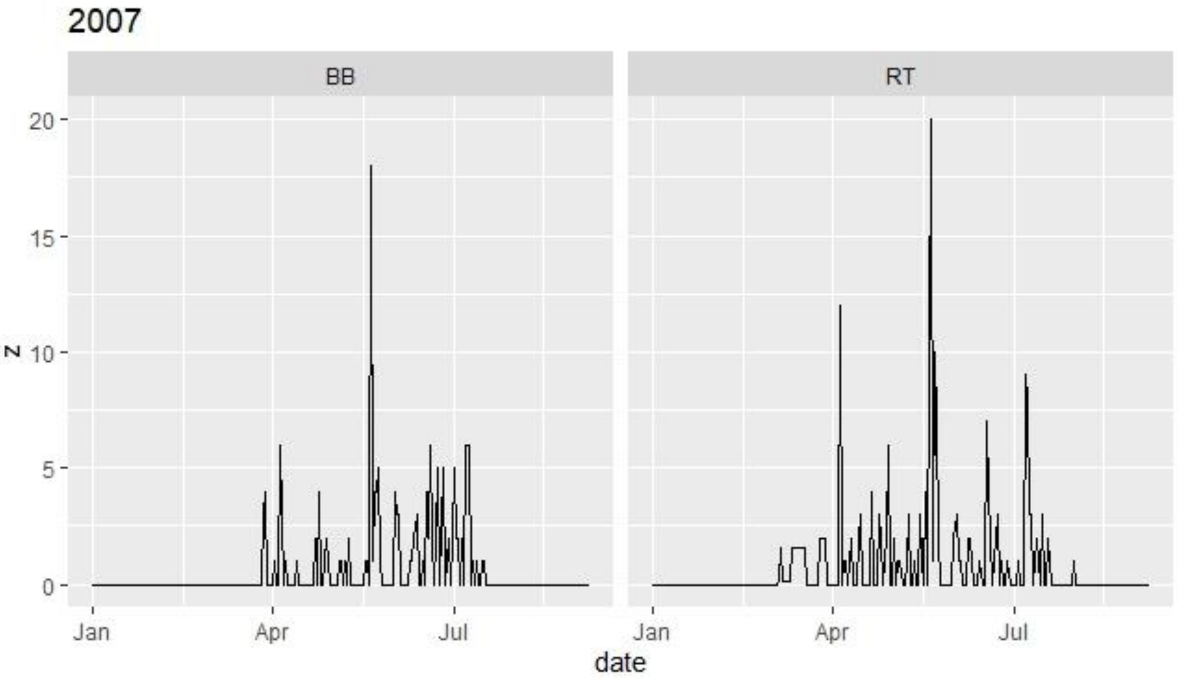

2002
BB
RT
z
40
20
0
Jan
Apr
Jul
Jan
Apr
Jul
date

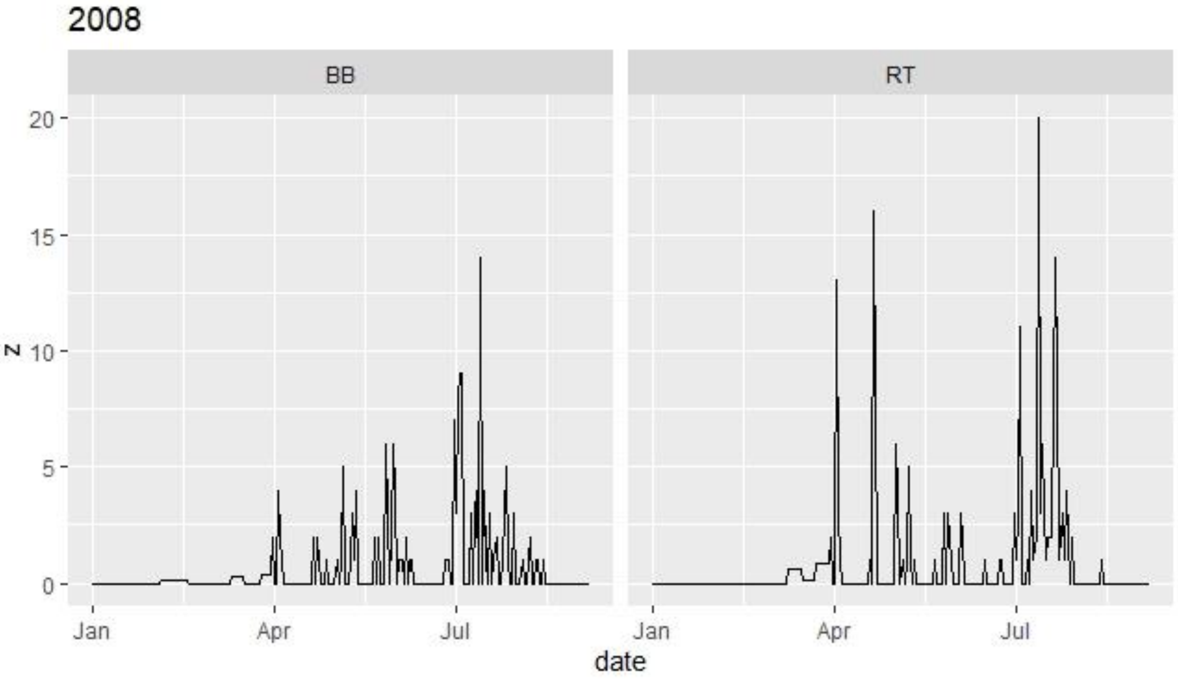

2003
BB
RT
z
20
10
0
Jan
Apr
Jul
Jan
Apr
Jul
date

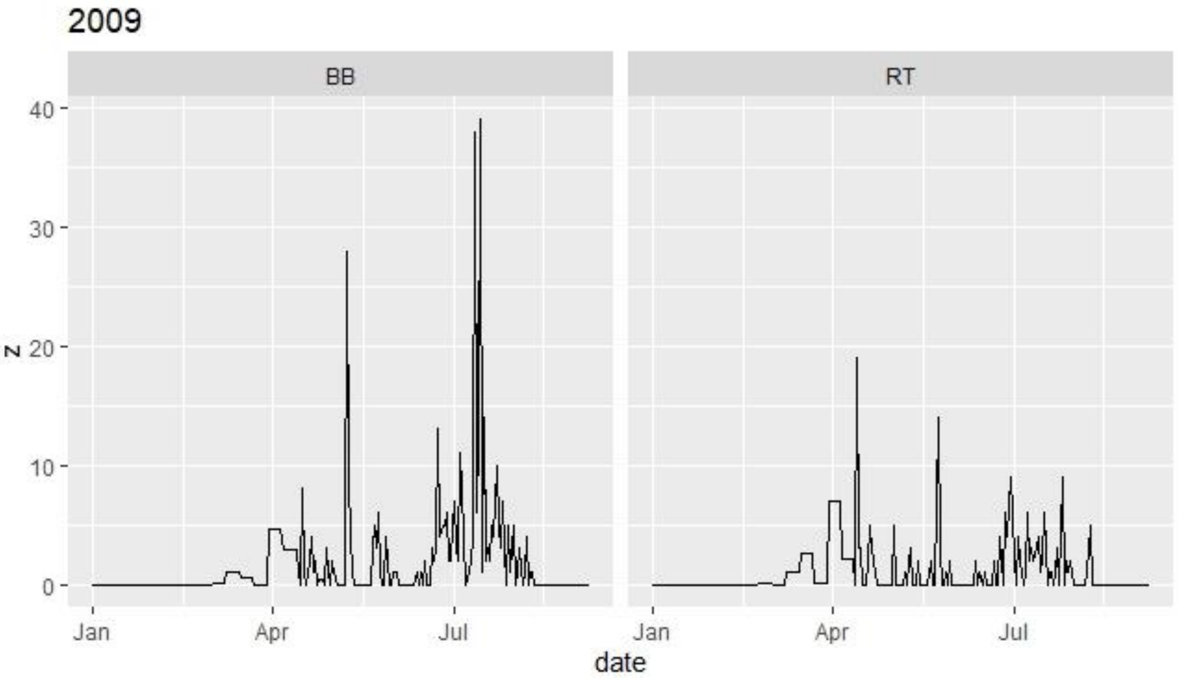

2004
BB
RT
z
120
80
40
0
Jan
Apr
Jul
Jan
Apr
Jul
date

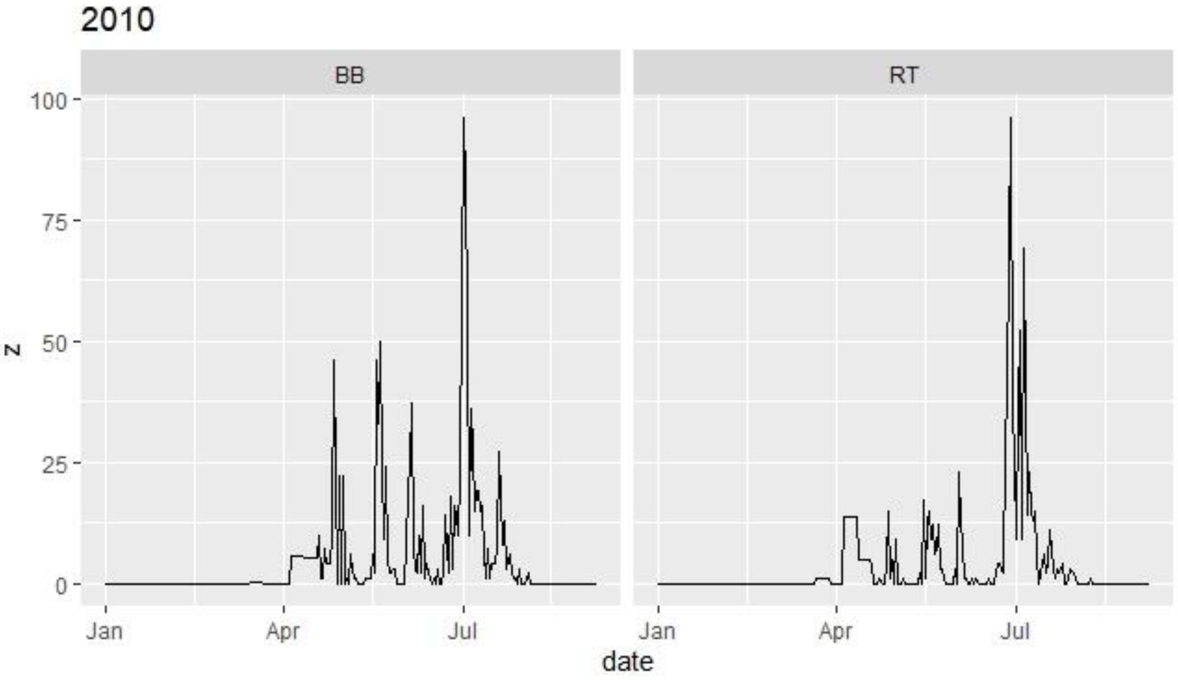

2005
BB
RT
z
0
5
10
15
Jan
Apr
Jul
Jan
Apr
Jul
date

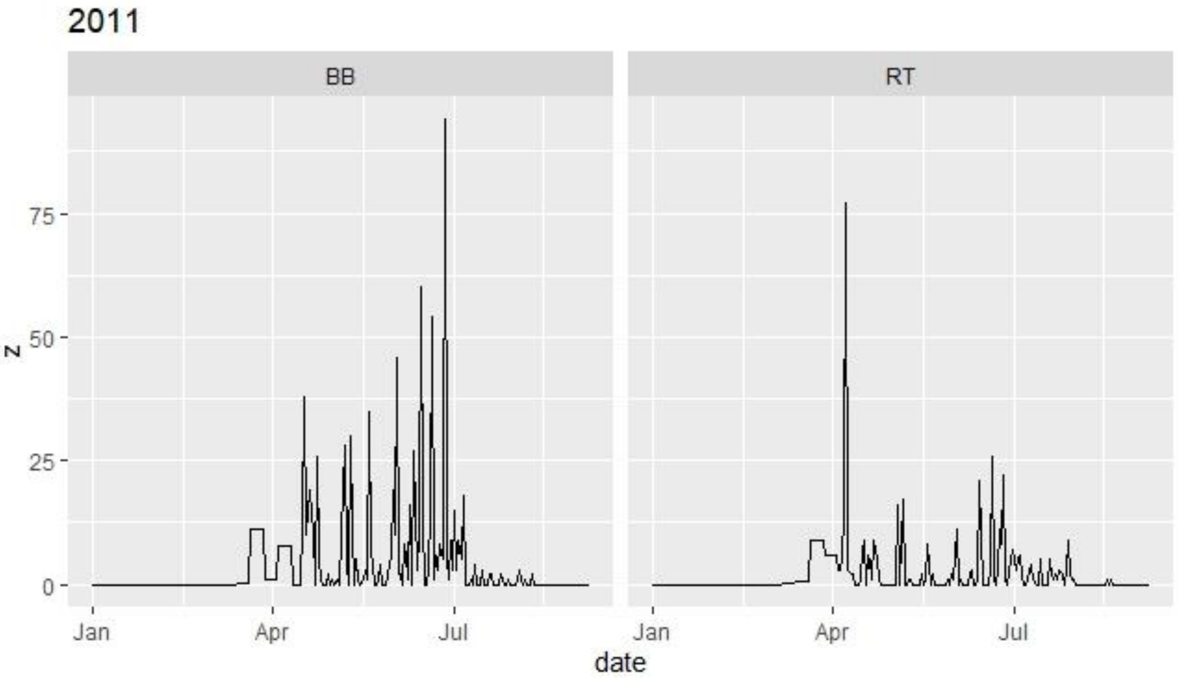

2006
BB
RT
z
0
5
10
15
Jan
Apr
Jul
Jan
Apr
Jul
date

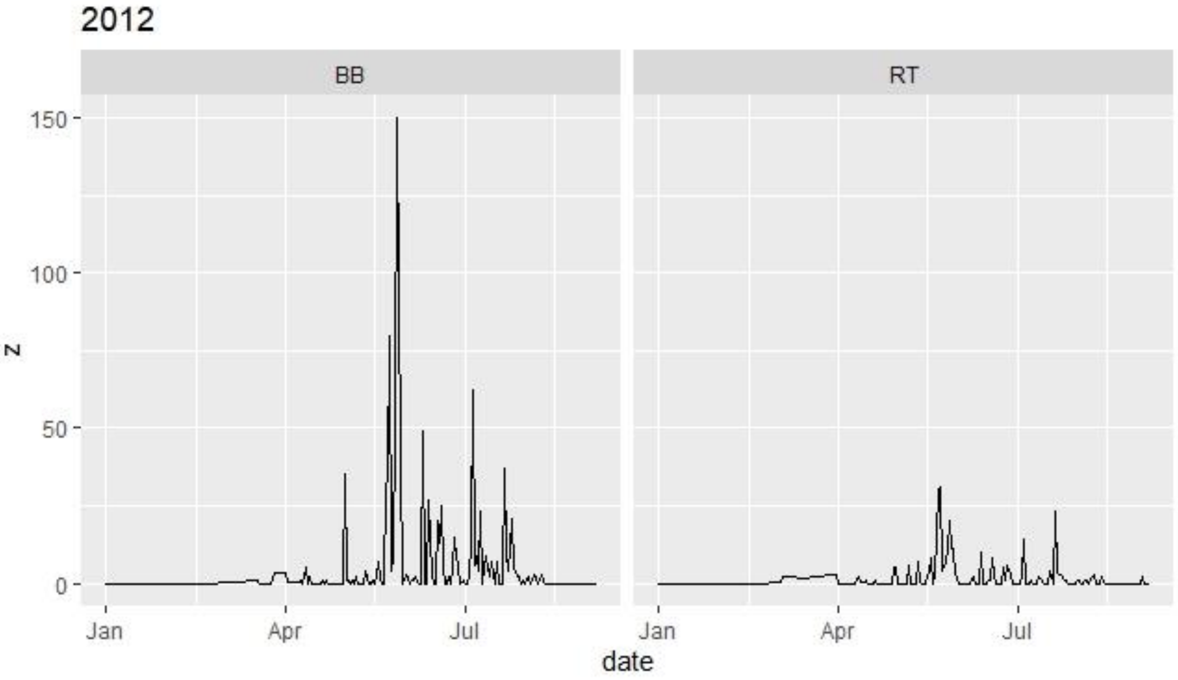

2007
BB
RT
z
0
5
10
15
20
Jan
Apr
Jul
Jan
Apr
Jul
date

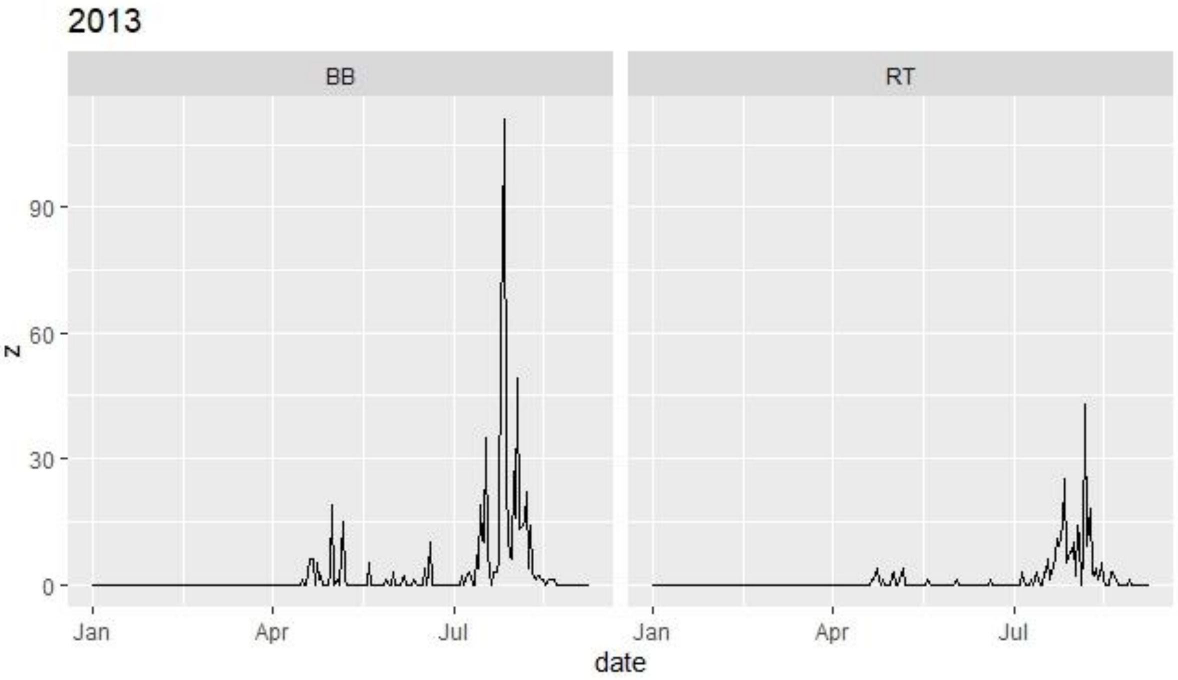

2008
BB
RT
z
0
5
10
15
20
Jan
Apr
Jul
Jan
Apr
Jul
date

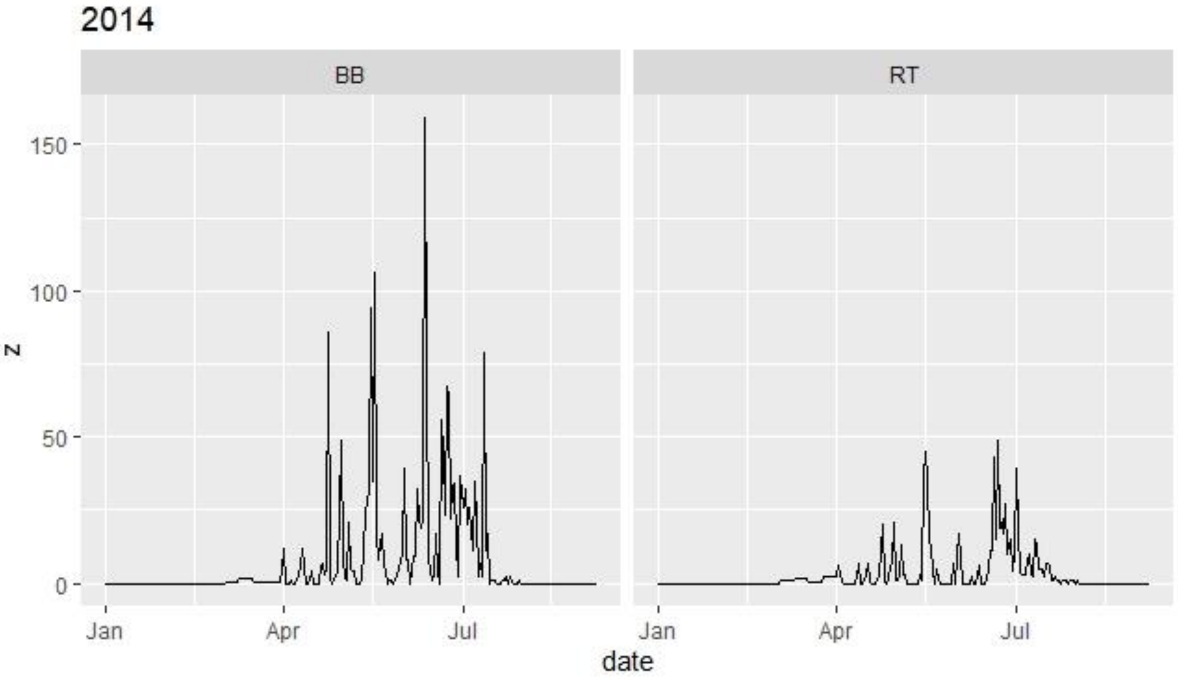

2009
BB
RT
z
0
10
20
30
40
Jan
Apr
Jul
Jan
Apr
Jul
date

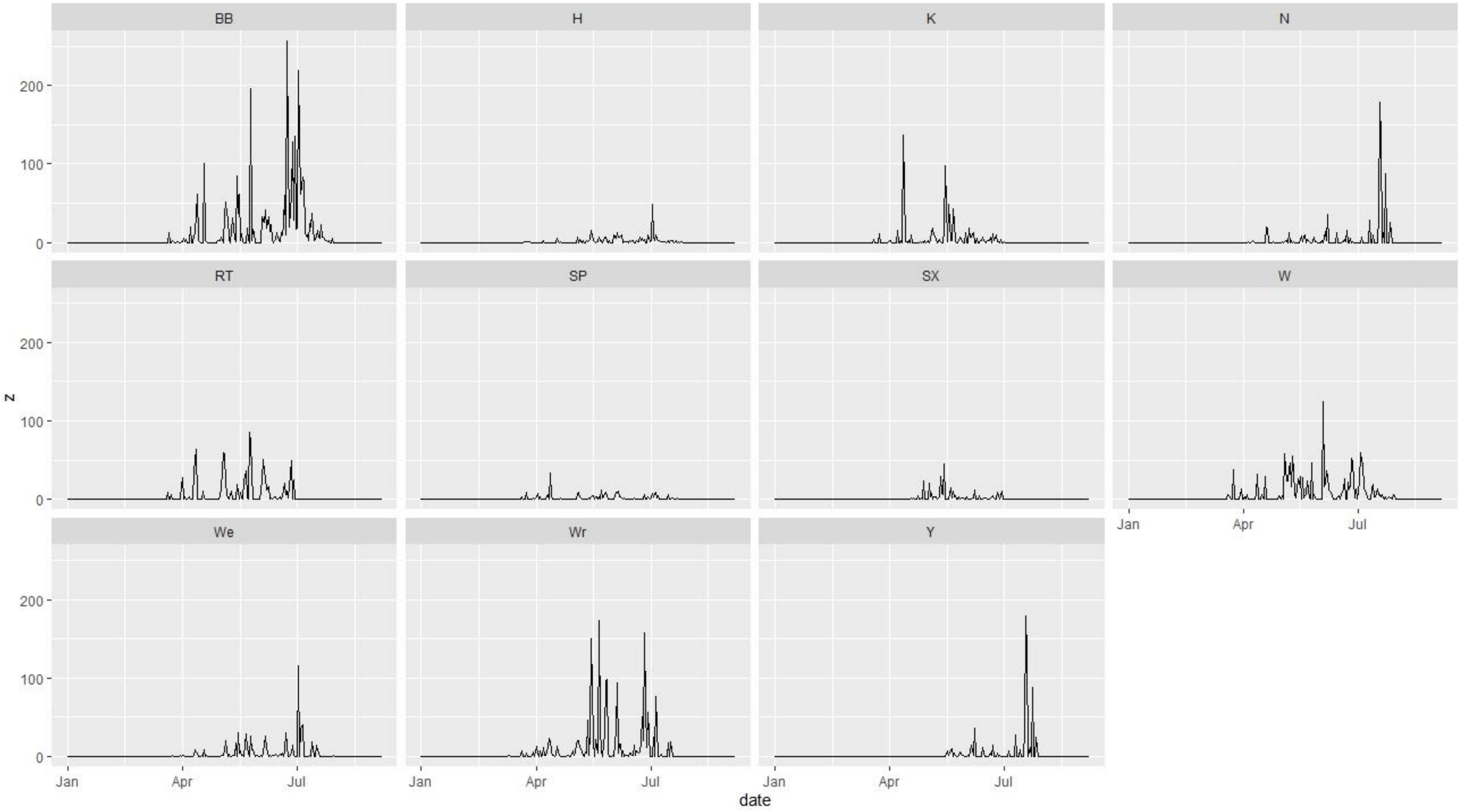

2010
BB
RT
z
0
25
50
75
100
Jan
Apr
Jul
Jan
Apr
Jul
date


2011
BB
RT
z
0
25
50
75
Jan
Apr
Jul
Jan
Apr
Jul
date

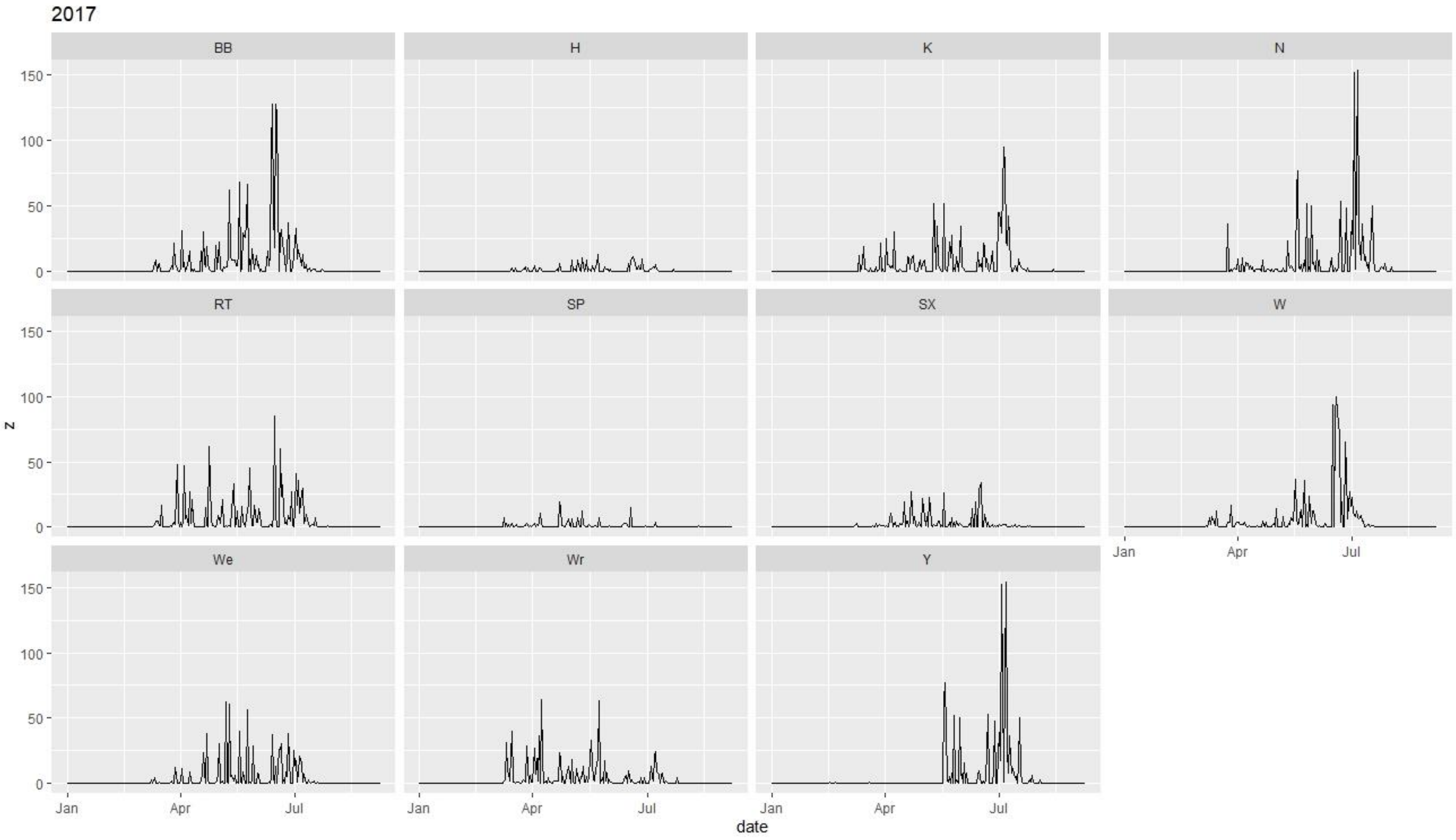

2012
BB
RT
z
0
50
100
150
Jan
Apr
Jul
Jan
Apr
Jul
date

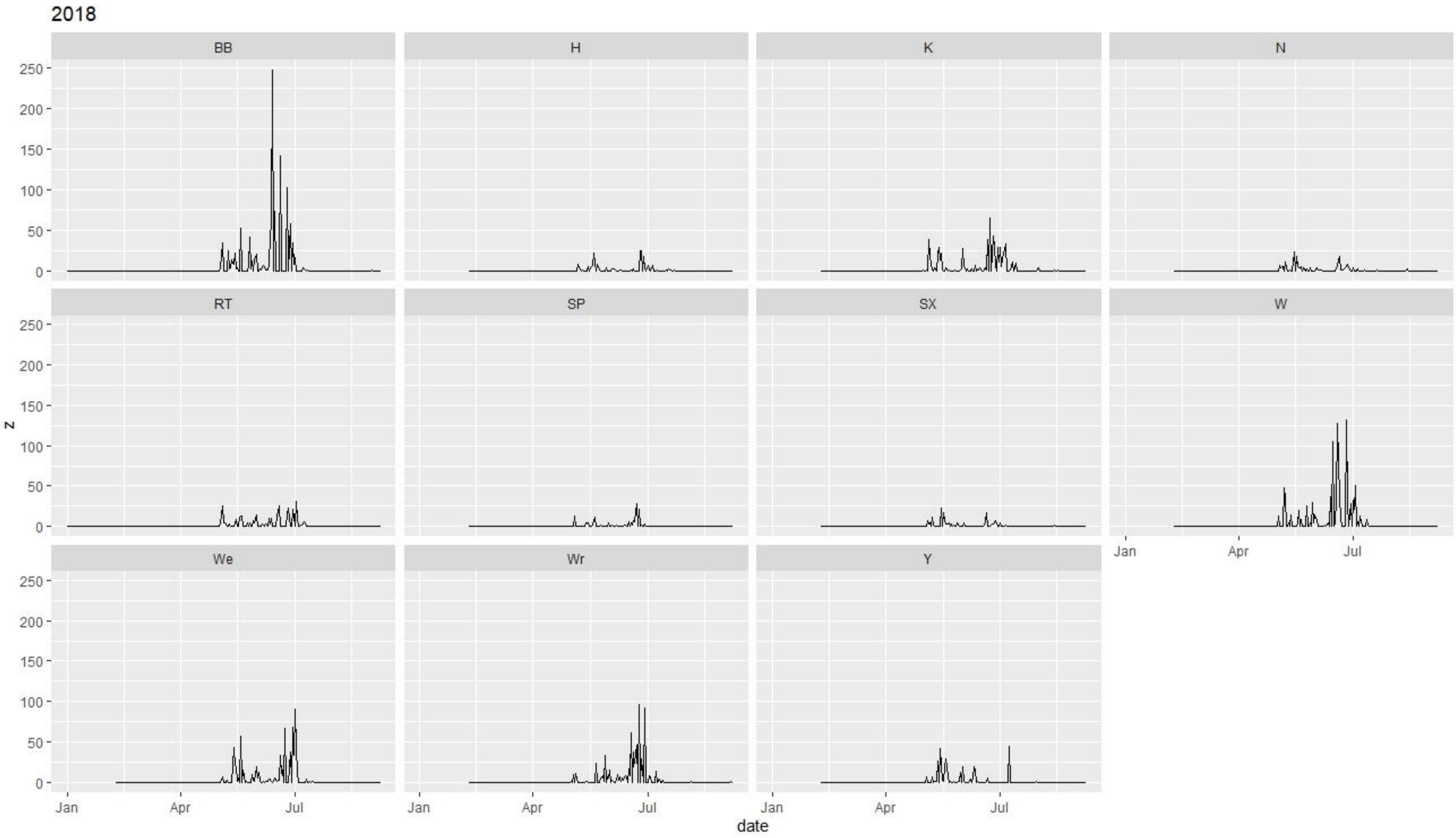

2013
BB
RT
z
90
60
30
0
Jan
Apr
Jul
Jan
Apr
Jul
date


2014
BB
RT
z
150
100
50
0
Jan
Apr
Jul
Jan
Apr
Jul
date



2016


2017
BB
H
K
N
RT
SP
SX
W
We
Wr
Y
z
0
50
100
150
Jan
Apr
Jul
date


2018
BB
H
K
N
RT
SP
SX
W
We
Wr
Y
z
250
200
150
100
50
0
Jan
Apr
Jul
date

Appendix C:

Study of pollen beetle eclosion using emergence traps within an oilseed rape crop

This study was initiated in order to gain a better understanding of how suction-trapped B. aeneus relate to newly-emerged Brassicogethes aeneus *in the field. The presence of teneral (not fully developed adult) beetles in the suction-traps indicates that they fly soon after eclosion. This work is still incomplete given the complications arising as described below*.

Abstract

Emergence traps have been used to assess the abundance of insects developing in the soil under crops but, given the potentially insulative material of the trap they have not generally been used to study phenology. Here I studied B. aeneus *eclosion (emergence from pupae) and compared their phenology and abundance to data from the suction-trap located on Rothamsted farm; I compared the temperature of the soil just below the surface inside emergence traps and outside within a mature oilseed rape crop on Rothamsted farm in two separate years. In year one,* B. *aeneus captures were recorded daily. In year two, pairs of temperature dataloggers were set in three different fields, one inside a trap and one outside. I found that emergence traps had a warming influence on the soil in the early summer but became a cooling influence later in the year.*

C.1 Introduction

Emergence traps (Southwood & Henderson 2000) have been used in entomology for many years to assess populations of insects emerging from the soil or low-growing vegetation (Richards &

Waloff, 1954, Jepson & Southwood 1958, Richards & Waloff 1961, Southwood & Jepson 1962a, 1962b Southwood & Siddorn 1965, Elliott & Powell 1966, Hadley 1969, Glen 1976, Way *et al*. 1969, Robinson 1979, Williamson *et al*. 1979, Lindelöw & Weslien 1986, Seddon 1986, Moeed & Meads 1987, Roubos & Liburd 2010) but, perhaps due to the warnings given by Southwood & Siddorn (1965) regarding the relative temperature under the trap only a few have been used in relation to insect phenology (Robinson 1979, Williamson *et al*. 1979, Lindelöw & Weslien 1986, Moeed & Meads 1987, Ferguson *et al*. 2003a, Ferguson *et al*. 2004, Roubos & Liburd 2010). More recently emergence traps have been used in field crops to study insect pest abundance in oilseed rape (*Brassica napus* L) (OSR) (Ferguson *et al*. 2003a, 2003b, 2006) using a specially designed emergence trap (Fig. C.1).

Southwood & Siddorn (1965) studied four different designs of emergence trap with respect to the temperature differential between inside and outside the trap. They found, for each trap design, that temperature differed between the trap and the control. To wit, cooler temperatures during the day with a slower loss of temperature at night in cloth-covered traps compared to metal traps (though both types remained warmer overnight than the control). The authors warned that these differences could have major effects on the development and phenology of the insects under the trap, however these issues were more important for assessing emergence of over-wintering populations than in the summer months. These conclusions were based on six weeks of recording during August in open habitat and thus potentially have restricted relevance outside that habitat and time period. Other authors have also assessed the effects of temperature on emergence traps (Hadley, 1969, Glen 1976) or on similar aphid cages (Woodford 1973). Hadley (1969) found similar differences – though the discrepancy was very small and was discounted. Glen (1976) found little difference but reported a dampening of temperature fluctuations and temperature differences were higher in less-shaded areas. Woodford (1973), working on much larger aphid-proof field cages in September, found that maximum temperatures were not significantly different during the day, however nocturnal temperatures were higher inside the cage than outside and temperature fluctuations were less inside the cage. The most substantial effects were that air movement was reduced by over 85% and light intensity reduced by up to 58%, though humidity remained similar inside and outside the cage (Woodford 1973) which may explain some of the temperature differences found in smaller, but otherwise similar emergence traps. Given these findings it is still unclear how much effect there is on the temperature under emergence traps when situated in tall vegetation – such as OSR.

C.2 Materials & Methods

C.2.1 Emergence Traps

The emergence traps (Fig. C.1) consist of a conical scaffold made from stainless steel over which a cloth mesh is fitted enclosing 0.5m^2 of ground (Ferguson *et al*. 2003b). The steel frame is made up of a 798mm circular base 600mm deep with three struts raising to a plastic top-piece 1.5m high. A collecting jar is attached to the head piece via a screw joint and consists of a central tube of 25mm diameter that opens into the jar which can be filled with trapping fluid (in this research this was made up of a solution of 60% industrial methylated spirits with 40% water). The trap base is buried in the soil to ~5cm depth to prevent egress & ingress from/into the trap. Emerging insects crawl or fly up to the head piece, pass through the opening and into the collecting jar. Traps are placed in the field in the late winter or early spring without the cloth cover. For sampling pollen beetles (*Brassicogethes aeneus* (Fabr.)). The crop is allowed to grow through the trap until green bud phase has passed, after which the crop is cut back in order to fit the mesh. The trap is then left for ~1 week to allow any adult insects under the trap to escape before the collecting jar is then put in place and checked daily for beetles.

Fig. C.1: Emergence trap with collecting net and collecting jar in place.

C.2.2 Temperature dataloggers

TinyTag Talk 2 – TK-4023 dataloggers (Gemini) were used, set to record the temperature every hour, on the hour and the minimum and maximum temperatures recorded during the hour. The dataloggers were buried just below the soil surface inside a sealable plastic bag (for added waterproofing), one inside the emergence trap and one ~1m away from the trap.

Fig. C.2: Field locations of the emergence traps. 2017 locations were Sawyers, Great Field and High Field, 2018 locations were Delharding, Little Knott and Pastures.

C.2.3 Study sites

Three fields were sampled (Fig. 3 a-c) in 2017 and 2018. In each field the emergence traps were placed in a rough square, with two traps ~5 metres from the field edge and two ~30-40 metres from the field edge.

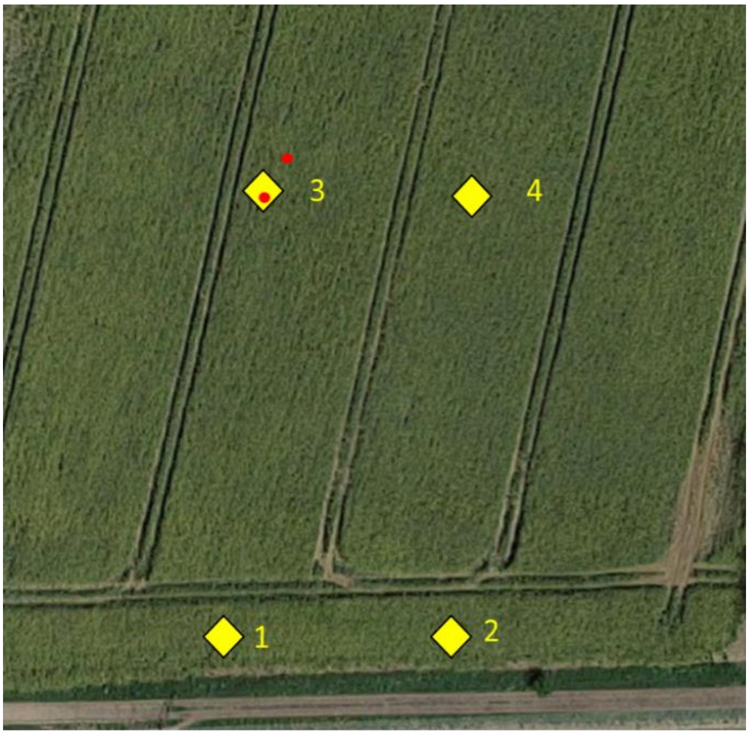


Fig. C.3: Schematic of sampling locations. The edge of the field is seen at the bottom of the picture, the yellow diamonds are the locations of emergence traps. The red dots show the locations of the TinyTag data loggers.

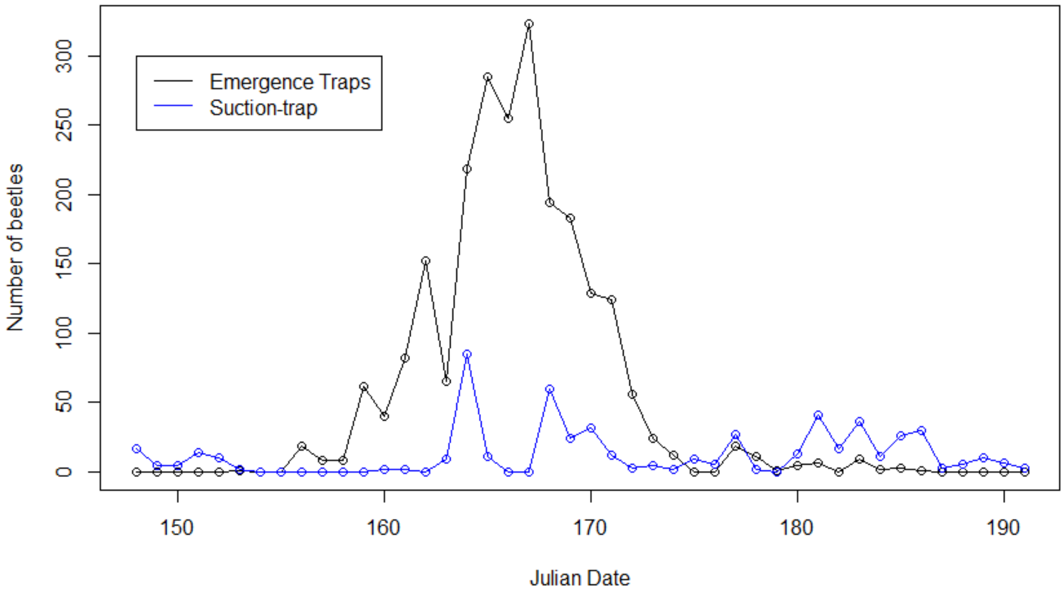
C.3 Results

C.3.1 Count data

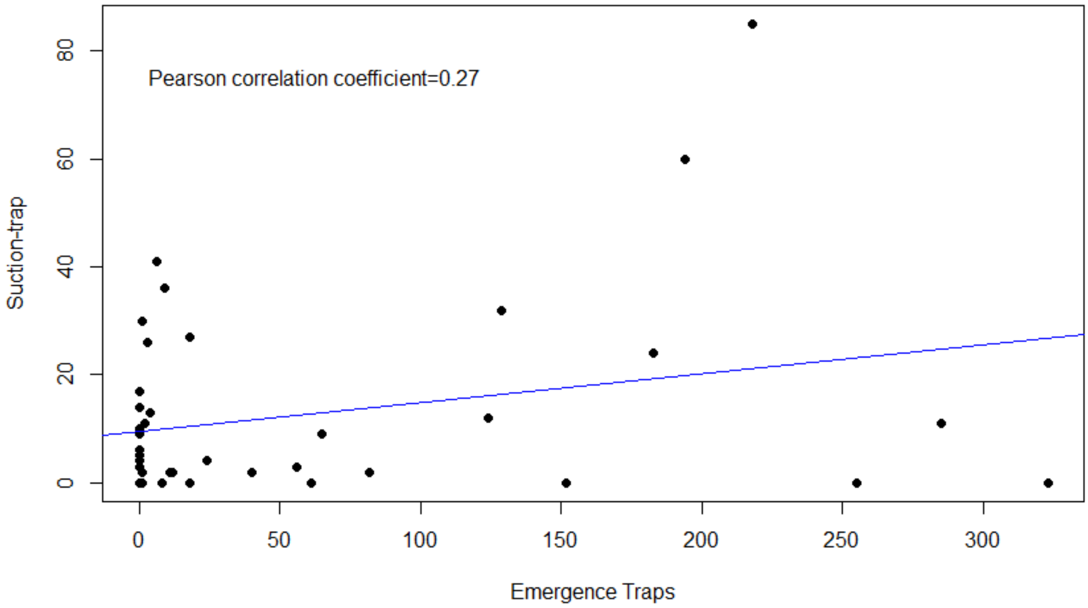
At time of writing, only 2017 data has been analysed. The count data show that emergence-trapped *B. aeneus* have a similar seasonality to suction-trapped beetles (Fig. C.4a), however the correlation between the two traps is poor (Fig. C.4b), with a Pearson correlation coefficient of 0.27. Interestingly, there is evidence that *B. aeneus* eclosion appears to lead local air temperature (Fig. C.4c), however this is likely an artefact of the effect of the emergence traps on the temperature of the soil beneath.

C.3.2 Temperature data

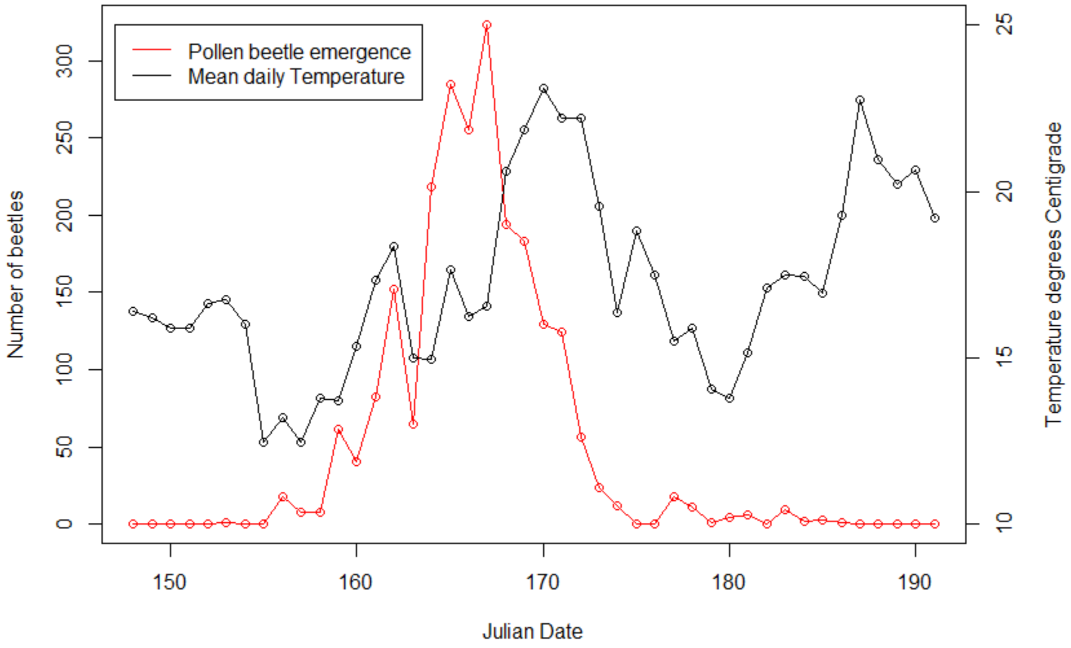
Mean daily temperature inside and outside the emergence traps were positively correlated (Fig. C.5) with Pearsons correlation coefficients of 0.99 (Delharding and Little Knott) and 0.98 (Pastures) returned. However, further investigation suggests this promising finding may not tell the full story. It appears that in the early part of the season the traps have a warming effect on the soil below them, with mean daily temperature inside the traps reaching more than a degree Centigrade warmer than outside at Delharding and Pastures (Fig. C.6). Later in the season, as air temperatures increase, the effect of the emergence trap on temperature appears to switch to a cooling one, with temperatures inside the trap more than a degree Centigrade cooler than outside at Pastures and nearly one degree difference at Delharding and Little Knott.


Emergence Traps
Suction-trap
Number of beetles
0
50
100
150
200
250
300
150
160
170
180
190
Julian Date

a


Pearson correlation coefficient=0.27
Suction-trap
0
20
40
60
80
0
50
100
150
200
250
300
Emergence Traps

b



c

Fig. C.4: Suction-trap and emergence trap data 2017. a: plot of daily counts of *Brassicogethes aeneus* at the Rothamsted suction-trap and in emergence traps (2017). b: Correlation plot of a. c: Daily emergence trap counts across all traps and air temperature recorded at Rothamsted meteorological station.

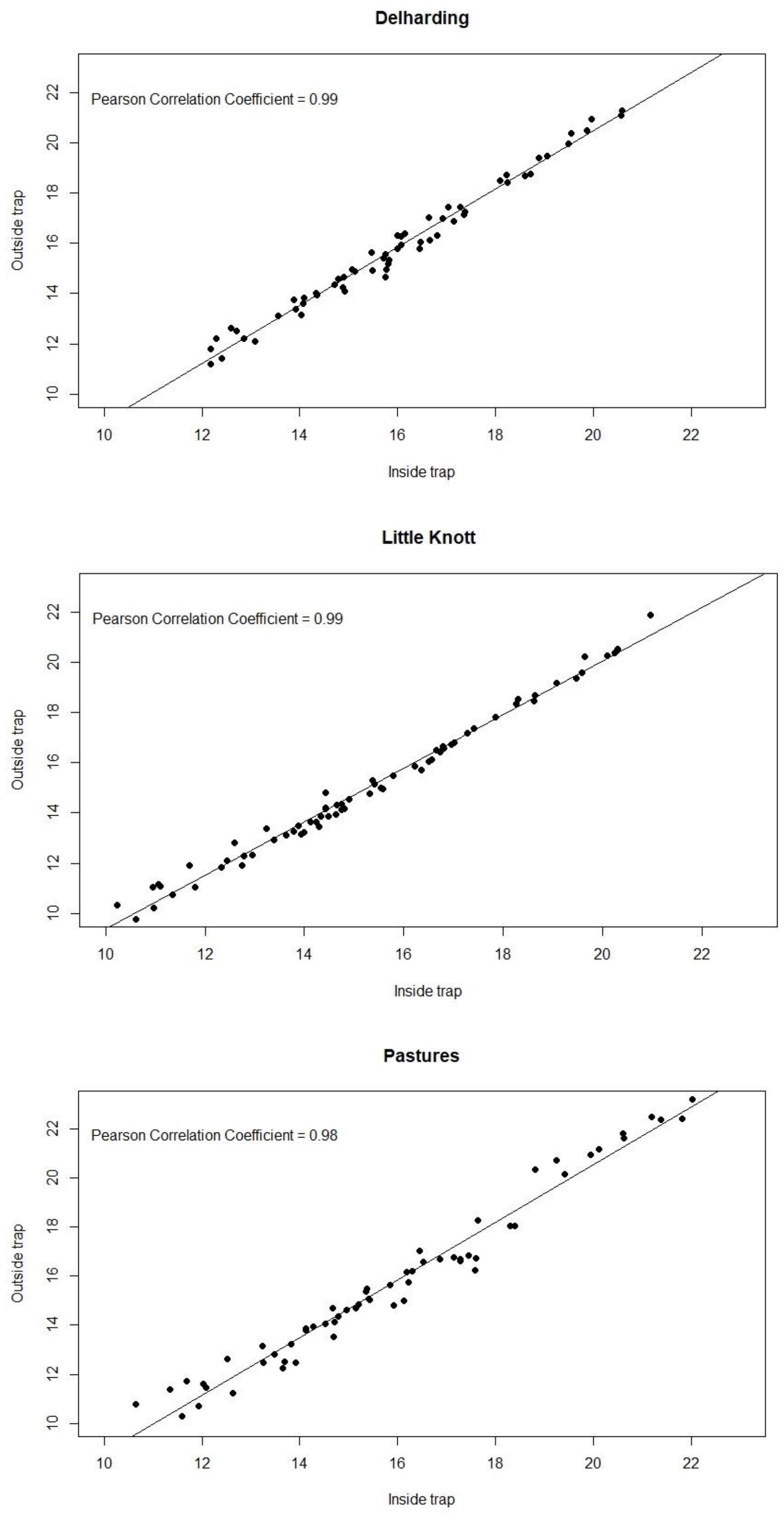


Fig. C.5: Correlation plots for TinyTag datalogger mean daily temperature inside and outside of emergence traps, with Pearson correlation coefficients given.

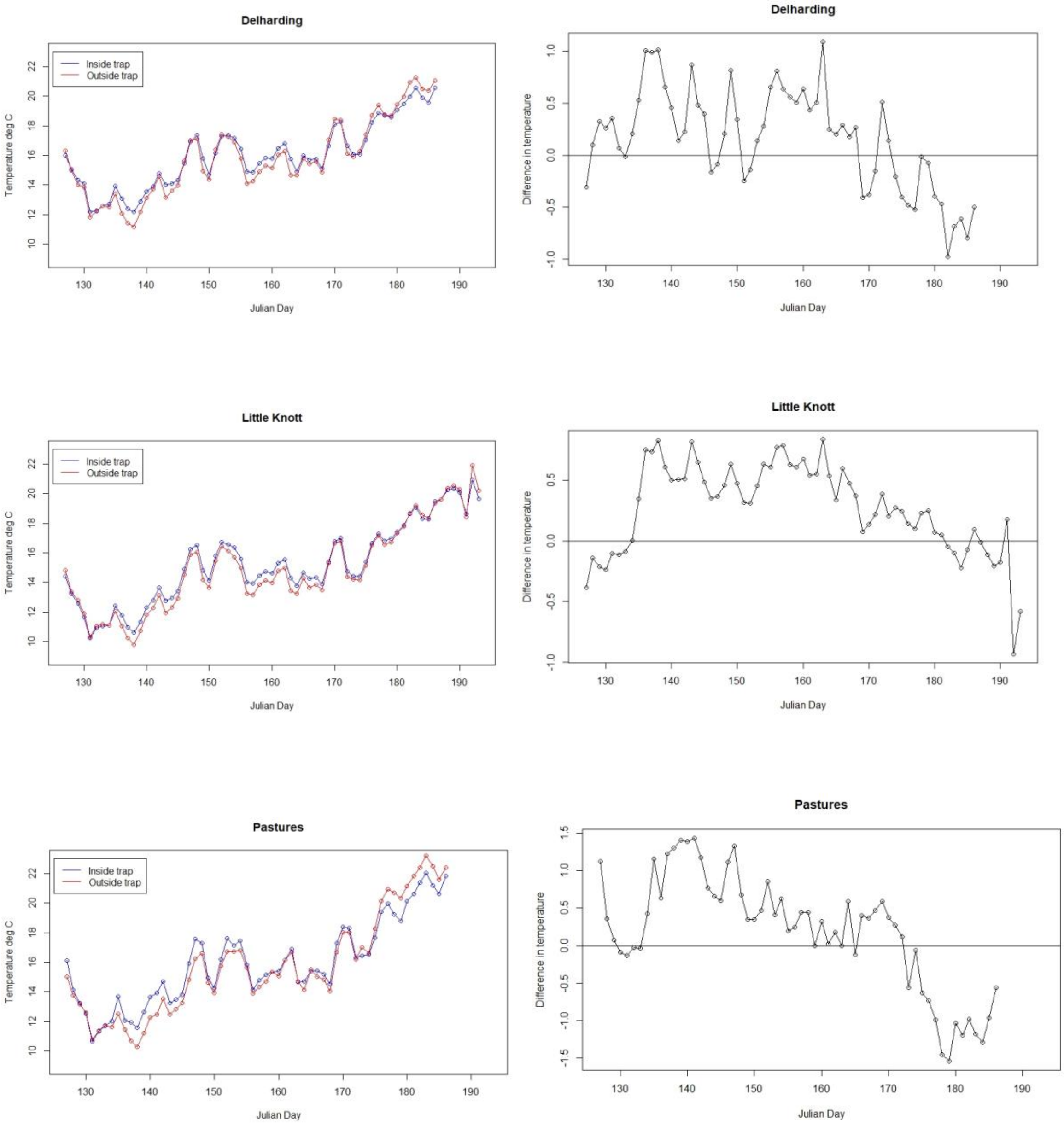


Fig. C.6: Mean daily temperature recorded (left) inside and outside emergence traps in three fields (top to bottom) sown to oilseed rape and the difference between these values (right).

C.4 Discussion

More work needs to be done on how the emergence trap affects temperature of the soil below and, in turn, how that affects the development of the pupating *B. aeneus.* The difference in temperature of the soil inside and outside of the emergence trap, whilst generally low (1.5°C at most) does have implications for using these traps to study the phenology of emergence. If it is the case that, despite the close correlation of temperatures inside and outside of the traps, the soil is warmed within the traps during May and June then this is likely to speed up pupal development (Bursell 1974, Cammell & Knight 1992) and may explain why eclosion appeared to lead temperature in 2017 (Fig. C.4c). This finding is complicated by the cooling effect two of the traps appear to have on the soil below from late June onwards (Fig. C.6), as this may, in turn, arrest pupal development compared to the wider environment. It should be noted that 2017 and 2018 are both in the ten hottest years on record in the UK and June 2018 was the driest June on record since 1925 (Kendon *et al*. 2019), which perhaps means that data from these years are non-typical. Questions also remain over whether this warming/cooling effect of the traps is restricted to those sited deep within the crop, or whether the effect is different closer to the field edge. A further experiment, not reported here was carried out in 2019 where traps with pairs of temperature loggers were set at different distances from the field edge. However, with some dataloggers failing there were an unsatisfactory number of replicates and the data are not yet fully analysed.

Appendix D: Collated Life cycle Parameters

Life cycle parameters were assembled from the literature with a view to using them in a modelling exercise using the Dymex software that proved unsuccessful. These parameters are presented here and may provide the basis for future efforts to produce a holistic (predictive) life cycle model. Parameters arising from empirical studies reported on by the paper are left unmarked. Parameters quoted from other papers are marked with a numerical superscript referring to the source (see final table). Parameters arising from modelling assumptions within the paper are marked with a superscript “x”, while unreferenced parameters reported within papers are marked with an asterisk “*”.

Spring Migration

	Start	Arrival in Crop
Blazejewska 1958	emerge at 9C	
Cook 2000	emerge at 10C[1] migrate at 15C[2]	
Ekbom & Ferdinand 2003	migrate when temp exceeds 14C	
Ferguson *et al.* 2015	Propensity to fly follows a sigmoid temperature–response curve in the 6–23 °C temperature range tested.	
Junk *et al*. 2016		8.0 °C; mean soil temperature 4.6 °C; and sunshine duration 3.4 h. The optimal cut-off for precipitation was 1.0 mm and optimal persistence of these conditions was one day only[x]
Nilsson 1988a	activity starts at 5-10C* long distance dispersal at 15C[2,3,4]	
Mauchline 2003	migrate at 12C* gregarious flights at 13.5C[5]	
Riggi *et al*. 2016	emerge at 11C*	
Seimandi Corda 2018	once temps reach 12C	
Skellern *et al.* 2017		10°C. Arrivals in crop increase until an accumulated temperature of 4.5 degree heating days, then plateau.
Stratanovich *et al*. 2014	when mean air temp exceeds 9C for 5 consecutive days[x]	
Tölle 2014	when air temp reaches 12C[2] fly long distances when temp reaches 15C*	
Williams 2010	emerge at 10C[1,6] migrate at 12C[7]	
Williams & Ferguson (2010)	12-15C*	

Egg

	Fecundity (eggs per female)	Maturation in Ovaries	Laying Density	Laying Rate	Egg Development	Mortality
Borg & Ekbom 1996				1131.6 seconds per egg		
Bromand 1983	17, 200-400[8] 67-344[9]	3.8 days		3.8 per female per day	25.6 days at 22C	9.7-15.7%
Cook 2000				2-3 per bud, up to 10[2,9,10,11] 1.39-2.12 eggs per bud		17-18%
Ekbom & Borg 1996	54.4 (restricted period)		0.125 eggs per bud (field), 0-0.91 eggs per bud (greenhouse)	2.9 eggs per female per plant, 4.22 eggs per 2 day period		
Ekbom & Ferdinand 2003	can be more than 200, average 54.4[10,11]	45-55 days		2.36 eggs per female per day[11]		
Hokkanen 2000	Only 40% of potential in field conditions					
Mauchline 2003	More than 200[12]	4-7 days*				
Nilsson 1988b				3.4-6.6 eggs per female per day		
Nilsson 1988c	95-185	7-10 days	Greater than 30,000/m2		45-50 days	Greenhouse: 6% (1-8%), Field estimation ~25%+6%
Scherney 1953	246	4-7 days at 16-21C				
Seimandi Corda 2018	Up to 200[9] up to 250*			Batches of 2-3 (occasionally 10)*		
Stratanovich *et al*. 2014	250 on average*			10 batches (minimum 15, maximum 35), every 85 day degrees[x]		33%*
Tölle 2014	78 (15-16C), 211 (20-22C), 206 (27C) all at 95% humidity[2] 200-300[13]	2-12 days[2,9,14,15]				
Veromann *et al*. 2011			~30 per plot (10 plants)			
Williams 2010	mean 246[9]					
Williams & Ferguson (2010)	80-180[11,16]					

Larva							
	Density	Development			Mortality		
		Overall	1st instar	2nd instar	Overall	1st instar	2nd instar
Bromand 1983	15.80%		3 days	3.1 days			5.80%
Buechi 2002					65.9-95.9%		
Cook 2000	0.19-0.96 larvae per 50 buds, 0.47-1.33 larvae per 20 flowers		4.23-6.11 days, 5-10 days[13]	5.3-5.87 days, 14 days[14]		37.4-67.1%	19.2-33.3%
Ekbom & Borg 1996	0.094 larvae per bud						
Hokkanen 2000		1 month*			survival of one/bud 60% of the time		
Mauchline 2003		9-13 days*					
Nilsson 1988c			5-10 days	17-25 days	50%[x]	5.3-29.3% (estimate)	
Riggi *et al*. 2017							
Scherney 1953		27-30 days at 16C					
Seimandi Corda 2018					88%[19]		
Skellern & Cook 2014		2 weeks*					
Skellern et al. 2017		2 weeks *					
Tölle 2014			2-10 days[2,9,14,15]	3-20 days[2,9,14,15]			
Veromann *et al*. 2011	15-250 per plot (10 plants)						
Williams 2010			5-10 days (Nilsson 1988c)	2 weeks[14]			
Williams & Ferguson (2010)			5-10 days*	14 days*			
Zaller *et al*. 2009	1797/m2						

Larval Parasitization

	Rate	Mortality
Billqvist 1998	Greenhouse: 0.048 (eggs & larvae), 0.039 (larvae), field: 0.039-0.057 (eggs & larvae) 0.038-0.040 (larvae)	
Hokkanen *et al*. 1988	21.8-49%, 50-90%[18]	
Nilsson 1988c	up to 51%	
Riggi *et al*. 2017		0.017% (0-0.24%)
Seimandi Corda 2018	67%, 16%[19], 97%[20]	
Tölle 2014	0% to more than 90%[15,21,22,23,24,25]	
Zaller *et al*. 2009	~0.48% (Positively correlated with abundance)	

Adult

	Start of Autumn Migration	Overwintering Mortality	General Mortality
Blazejewska 1958	"middle of August, disregarding the temperature"		
Hokkanen 1993		85-98%	
Seimandi Corda 2018			5.41 (♀) & 6.66 (♂) days (lab conditions inc. starvation)
Stratanovich *et al*. 2014	when mean air temp falls below 12C for 5 days*		

List of sources for the above tables:

1: Láska & Kocourek 1991	6: Nilsson 1988a	11: Ekbom & Borg 1996	16: Blazejewska 1958	21: Nilsson & Andreasson 1987
2: Fritsche 1957	7: Free & Williams 1978	12: Hopkins & Ekbom 1996	17: Borg & Ekbom 1996	22: Hokkanen 2000
3: Müller 1941	8: Börner & Blunk 1920	13: Nilsson 1988c	18: Jourdhueil 1961	23: Büchi 2002
4: Masurat 1966	9: Scherney 1953	14: Bromand 1983	19: Riggi *et al*. 2017	24: Williams 2006
5: Šedivý & Kocourek 1994	10: Nilsson 1988b	15: Nielsen & Axelsen 1988a	20: Ulber *et al*. 2006	25: Thies *et al*. 2008

References

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G. S., Davis, A., Dean, J., Devin, M. & Ghemawat, S. (2016). Tensorflow: Large-scale machine learning on heterogeneous distributed systems. *arXiv preprint arXiv:1603.04467*. http://tensorflow.org

Abd El-Ghany, N. M., Abd El-Aziz, S. E., & Marei, S. S. (2020). A review: application of remote sensing as a promising strategy for insect pests and diseases management. *Environmental Science and Pollution Research*, **27**, 33503–33515.

AHDB (2013). *Monitoring and control of pollen beetle in oilseed rape*. Information Sheet 18/Spring 2013.

Agrawal, R. & Mehta, S. C. (2007). Weather based forecasting of crop yields, pests and diseases – IASRI models. *Journal of the Indian Society of Agricultural Statistics*, **61**, 255-263.

Ahmad, M. N., Shariff, A. R. M., & Moslim, R. (2018). Monitoring insect pest infestation via different spectroscopic techniques. *Applied Spectroscopy Reviews*, **53**, 836-853.

Ahmed, O. S., Franklin, S. E., Wulder, M. A., & White, J. C. (2015). Characterizing stand-level forest canopy cover and height using Landsat time series, samples of airborne LiDAR, and the Random Forest algorithm. *ISPRS Journal of Photogrammetry and Remote Sensing*, **101**, 89-101.

Akaike, H. (1973). Information theory and the maximum likelihood principle. In: *2nd International Symposium on Information Theory*. Akademiai Ki à do, Budapest.

Alford, D. V., Carden, P. W., Dennis, E. B., Gould, H. J., & Vernon, J. D. R. (1979). Monitoring codling and tortrix moths in United Kingdom apple orchards using pheromone traps. *Annals of Applied Biology*, ***91***, 165-178.

Alford, D. V., Nilsson, C., & Ulber, B. (2003). Insect pests of oilseed rape crops. In: *Biocontrol of oilseed rape pests*, 9-41.

Allender, C. J., & King, G. J. (2010). Origins of the amphiploid species *Brassica napus* L. investigated by chloroplast and nuclear molecular markers. *BMC plant biology*, **10**, 54.

Allendorf, F. W. (1986). Genetic drift and the loss of alleles versus heterozygosity. *Zoo Biology*, **5**, 181-190.

Almond, J. A., Dawkins, T. C. K., & Askew, M. F. (1986). Aspects of crop husbandry. In: *Oilseed rape* (pp 127-175).

Appelqvist, L. A., Ohlson, R., & Sprague, M. A. (1973). Rapeseed: Cultivation, Composition, Processing and Utilization. *Soil Science*, **116**, 453.

Armstrong, J. S. (2001). Selecting forecasting methods. In: *Principles of forecasting* (pp. 365-386). Springer, Boston, MA.

Atakan, E., & Canhilal, R. (2004). Evaluation of yellow sticky traps at various heights for monitoring cotton insect pests. *Journal of Agricultural and Urban Entomology*, **21**, 15-24.

Audisio, P., Cline, A. R., De Biase, A., Antonini, G., Mancini, E., Trizzino, M., Costantini, L., Strika, S., Lamanna, F. & Cerretti, P. (2009). Preliminary re-examination of genus-level taxonomy of the pollen beetle subfamily Meligethinae (Coleoptera: Nitidulidae). *Acta Entomologica Musei Nationalis Pragae*, **49**, 341-504.

Audisio, P., Sabatelli, S. & Jelínek, J. (2014). Revision of the pollen beetle genus *Meligethes* (Coleoptera: Nitidulae). *Fragmenta entomologica,* **46**, 19-112.

Aukema, B. H., Carroll, A. L., Zhu, J., Raffa, K. F., Sickley, T. A., & Taylor, S. W. (2006). Landscape level analysis of mountain pine beetle in British Columbia, Canada: spatiotemporal development and spatial synchrony within the present outbreak. *Ecography*, **29**, 427-441.

Austin, C. J., Guglielmo, C. G., & Moehring, A. J. (2014). A direct test of the effects of changing atmospheric pressure on the mating behavior of Drosophila melanogaster. *Evolutionary Ecology*, **28**, 535-544.

Austin, M. P. (2002). Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. *Ecological Modelling*, **157**, 101-118.

Balaso, G. M., Nalini, C., Yankit, P., & Thakur, P. (2019). Push-pull strategy: Novel approach of pest management. *Journal of Entomology and Zoology Studies* **7**: 220-223.

Bajwa, W. I., Coop, L., & Kogan, M. (2003). Integrated pest management (IPM) and Internet-based information delivery systems. *Neotropical Entomology*, **32**, 373-383.

Baldock, K. C. R., Goddard, M. A., Hicks, D. M., Kunin, W. E., Mitschunas, N., Osgathorpe, L. M., Potts, S. G., Robertson, K. M., Scott, A. V., Stone, G. N., Vaughan, I. P. & Memmott, J. (2015) Where is the UK's pollinator biodiversity? The importance of urban areas for flower-visiting insects. *Proceedings of the Royal Society B: Biological Sciences*, **282**, 20142849.

Barbu, C., Dumonteil, E., & Gourbiere, S. (2010). Characterization of the dispersal of non-domiciliated *Triatoma dimidiata* through the selection of spatially explicit models. *PLoS Neglected Tropical Diseases*, **4**, e777.

Bardner, R. French, R. A. & Dupuch, M. (1982) Agricultural Benefits of the Rothamsted Aphid Bulletin, *Annual Report Rothamsted for 1981*.

Bardner, R., Fletcher, K. E., Griffiths, D. C., & Hamon, N. (1982). *Sitona lineatus* L. on faba beans in Britain. *FABIS Newsletter (ICARDA)*.

Barkan, S., Hoffman, A., Hezroni, A., & Soroker, V. (2018). Flight performance and dispersal potential of red palm weevil estimated by repeated flights on flight mill. *Journal of insect Behavior*, **31**, 66-82.

Barry, S. C., & Welsh, A. H. (2002). Generalized additive modelling and zero inflated count data. *Ecological Modelling*, **157**, 179-188.

Bartlet, E., Blight, M. M., Pickett, J. A., Smart, L. E., Turner, G., & Woodcock, C. M. (2004). Orientation and feeding responses of the pollen beetle, *Meligethes aeneus*, to candytuft, *Iberis amara*. *Journal of chemical ecology*, **30,** 913-925.

Bauer, S., Chapman, J. W., Reynolds, D. R., Alves, J. A., Dokter, A. M., Menz, M. M., Sapir, N., Ciach, M., Pettersson, L. B., Kelly, J. F. & Leijnse, H. (2017). From agricultural benefits to aviation safety: realizing the potential of continent-wide radar networks. *BioScience*, **67**, 912-918.

Bell, J. R., Alderson, L., Izera, D., Kruger, T., Parker, S., Pickup, J., Shortall, C. R., Taylor, M. S., Verrier, P. & Harrington, R. (2015). Long-term phenological trends, species accumulation rates, aphid traits and climate: five decades of change in migrating aphids. *Journal of Animal Ecology,* **84.** 21-34.

Bell, J. R., Blumgart, D., & Shortall, C. R. (2020). Are insects declining and at what rate? An analysis of standardised, systematic catches of aphid and moth abundances across Great Britain. *Insect Conservation and Diversity*, **13**, 115-126.

Bell, J. R., Botham, M. S., Henrys, P. A., Leech, D. I., Pearce-Higgins, J. W., Shortall, C. R., Brereton, T. M., Pickup, J. & Thackeray, S. J. (2019). Spatial and habitat variation in aphid, butterfly, moth and bird phenologies over the last half century. *Global Change Biology*, **25**, 1982-1994.

Bell, J. R., Clark, S. J. *et al.* (2021) Neonicotinoids disrupt spatial synchrony (Manuscript in preparation, title to be confirmed).

Bell, J. R., Taylor, M. S., Shortall, C. R., Welham, S. J., & Harrington, R. (2012). The trait and host plant ecology of aphids and their distribution and abundance in the United Kingdom. *Global Ecology and Biogeography*, **21**, 405-415.

Bengio, Y. (2012). Practical recommendations for gradient-based training of deep architectures. In: *Neural networks: Tricks of the trade* (pp. 437-478). Springer, Berlin, Heidelberg.

Bennett, E. J., Brignell, C. J., Carion, P. W., Cook, S. M., Eastmond, P. J., Teakle, G. R., Hammond, J. P., Love, C., King, G. J., Roberts, J. A. & Wagstaff, C. (2017). Development of a statistical crop model to explain the relationship between seed yield and phenotypic diversity within the *Brassica napus* genepool. *Agronomy*, **7**, 31.

Bennetts, R. E., Nichols, J. D., Pradel, R., Lebreton, J. D., & Kitchens, W. M. (2001). Methods for estimating dispersal probabilities and related parameters using marked animals. In: *Dispersal* (pp 3-17) Oxford University Press.

Berry, P. M., & Spink, J. H. (2006). A physiological analysis of oilseed rape yields: past and future. *The Journal of Agricultural Science*, **144**, 381.

Biesmeijer, J. C., Roberts, S. P., Reemer, M., Ohlemüller, R., Edwards, M., Peeters, T., Schaffers, A. P., Potts, S. G., Kleukers, R. J. M. C., Thomas, C. D. & Settele, J. (2006). Parallel declines in pollinators and insect-pollinated plants in Britain and the Netherlands. *Science*, **313**, 351-354.

Billqvist, A. (1998). The effect of host plant on parasitism of the pollen beetle (*Meligethes spp*.): a comparison between spring rape (*Brassica napus*) and white mustard (*Sinapis alba). Examensarbete i Entomologi (Sweden)*.

Billqvist, A., & Ekbom, B. (2001). The influence of host plant species on parasitism of pollen beetles (*Meligethes* spp.) by *Phradis morionellus*. *Entomologia Experimentalis et Applicata*, **98**, 41-47.

Bjørnstad, O. N., & Bjørnstad, M. O. N. (2016). Package ‘ncf’. *Spatial Nonparametric Covariance Functions,* 1-42.

Bjørnstad, O. N., & Falck, W. (2001). Nonparametric spatial covariance functions: estimation and testing. *Environmental and Ecological Statistics*, **8**, 53-70.

Bjørnstad, O. N., & Grenfell, B. T. (2001). Noisy clockwork: time series analysis of population fluctuations in animals. *Science*, **293**, 638-643.

Bjørnstad, O. N., Ims, R. A., & Lambin, X. (1999). Spatial population dynamics: analyzing patterns and processes of population synchrony. *Trends in Ecology & Evolution*, **14**, 427-432.

Blazejewska, A. (1958). *Zimowanie slodyszka rzepakowego (*Meligethes aeneus *Fabr.)*. Panstw. Wydawn. Naukowe.

Blight, M. M., Pickett, J. A., Ryan, J., Wadhams, L. J., & Woodcock, C. M. (1995). Recognition of oilseed rape volatiles by pollen beetles, *Meligethes* spp.: Electrophysiological and chemical studies. In: *Proceedings, GCIRC 9th international rapeseed congress*, **3**, 1043-1045.

Blight, M. M., & Smart, L. E. (1999). Influence of visual cues and isothiocyanate lures on capture of the pollen beetle, *Meligethes aeneus* in field traps. *Journal of Chemical Ecology*, **25**, 1501-1516.

Boerner, C., Blunck, H., & Dyckerhoff, H. (1921). Versuche zur Bekämpfung der Kohlerdflöhe und Rapsglanzkäfer. *Mittlg Biol Reichsanstalt Berlin*, **22**, 5-41.

Boiteau, G., & Colpitts, B. (2004). The potential of portable harmonic radar technology for the tracking of beneficial insects. *International Journal of Pest Management*, **50**, 233-242.

Booth, E. J., & Gunstone, F. D. (2004) Rapeseeds and rapeseed oil: agronomy, production, and trade. In: *Rapeseed and Canola Oil: Production, Processing, Properties and Uses.* ed. Gunstone, F. D. CDC Press, USA.

Borden, J. H., & Bennett, R. B. (1969). A Continously Recording Flight Mill for Investigating the Effect of Volatile Substances on the Flight of Tethered Insects. *Journal of Economic Entomology*, **62**, 782-785.

Borg, A., & Ekbom, B. (1996). Characteristics of oviposition behaviour of the pollen beetle, *Meligethes aeneus* on four different host plants. *Entomologia experimentalis et applicata*, **81**, 277-284.

Börner, C. & Blunck, H. (1920). On the life history of *Meligethes aeneus* F. *Mitt. biol. ZentAnst., Berl.*, no. 18, 15th Yearly Rep. pp. 91–109.

Boulanger, Y., Fabry, F., Kilambi, A., Pureswaran, D. S., Sturtevant, B. R., & Saint-Amant, R. (2017). The use of weather surveillance radar and high-resolution three dimensional weather data to monitor a spruce budworm mass exodus flight. *Agricultural and Forest Meteorology*, **234**, 127-135.

Bourhis, Y., Bell, J. R., van den Bosch, F., & Milne, A. E. (2020). Artificial neural networks for monitoring network optimisation—a practical example using a national insect survey. *Environmental Modelling & Software*, 104925.

Bradter, U., Kunin, W. E., Altringham, J. D., Thom, T. J., & Benton, T. G. (2013). Identifying appropriate spatial scales of predictors in species distribution models with the random forest algorithm. *Methods in Ecology and Evolution*, **4**, 167-174.

Brattsten, L. B., Holyoke, C. W., Leeper, J. R., & Raffa, K. F. (1986). Insecticide resistance: challenge to pest management and basic research. *Science*, **231**, 1255-1260.

Breiman, L. (2001). Random forests. *Machine Learning*, **45**, 5-32.

Breiman, L., Friedman, J. H., Olshen, R. A., and Stone, C. G. (1984). *Classification and Regression Trees*. Wadsworth International Group, Belmont, California, USA.

Broadbent, L. (1948). Aphis migration and the efficiency of the trapping method. *Annals of Applied Biology*, **35**, 379-394.

Brodick, P. G., Davies, A. B. & Asner, G. P. (2019). Uncovering ecological patterns with convolutional neural networks. *Trends in Ecology and Evolution* **34**, 734-745.

Bromand, B. (1983). Possibility of continuous rearing of *Meligethes aeneus* Fabr.(Col.)[rape, *Brassica napus*]. *Zeitschrift fuer Angewandte Entomologie (Germany, FR)*.

Broquet, T., & Petit, E. J. (2009). Molecular estimation of dispersal for ecology and population genetics. *Annual Review of Ecology, Evolution, and Systematics*, **40**, 193-216.

Brouwers, N. C., & Newton, A. C. (2009). Movement rates of woodland invertebrates: a systematic review of empirical evidence. *Insect Conservation and Diversity*, **2**, 10-22.

Brown, J. K., Beeby, R., & Penfield, S. (2019). Yield instability of winter oilseed rape modulated by early winter temperature. *Scientific Reports*, **9**, 1-9.

Brown, T. M. (1981). Countermeasures for insecticide resistance. *Bulletin of the ESA*, **27**, 198-201.

Bruce, T. J., Wadhams, L. J., & Woodcock, C. M. (2005). Insect host location: a volatile situation. *Trends in Plant Science*, **10**, 269-274.

Bruce, T. J., Hooper, A. M., Ireland, L., Jones, O. T., Martin, J. L., Smart, L. E., Oakley, J. & Wadhams, L. J. (2007). Development of a pheromone trap monitoring system for orange wheat blossom midge, *Sitodiplosis mosellana*, in the UK. *Pest Management Science: formerly Pesticide Science*, **63**, 49-56.

Bryant, S. R. & Shreeve, T. G. (2002). The use of artificial neural networks in ecological analysis: estimating microhabitat temperature. *Ecological Entomology* **27**, 424-432.

Büchi, R. (2002). Mortality of pollen beetle (*Meligethes* spp.) larvae due to predators and parasitoids in rape fields and the effect of conservation strips. *Agriculture, Ecosystems & Environment*, **90**, 255-263.

Büchs, W., & Alford, D. V. (2003). Predators of oilseed rape pests. In: *Biocontrol of oilseed rape pests*, 181-200.

Büchs, W., & Nuss, H. (2000). First steps to assess the importance of epigaeic active polyphagous predators on oilseed rape insect pests with soil pupating larvae. *IOBC-WPRS Bulletin*, **23**, 151-164.

Buckner, C. H. (1966). The role of vertebrate predators in the biological control of forest insects. *Annual Review of Entomology*, **11**, 449-470.

Buechi, R. (1990). Investigations on the use of turnip rape as trap plant to control oilseed rape pests. *IOBC-WPRS Bulletin*, **13**, 32-39.

Bunting, E. S. (1984) Oilseed rape in perspective: with particular reference to crop expansion and distribution in England, 1973-1983. *Aspects of Applied Biology 6: Agronomy, physiology, plant breeding and crop protection of oilseed rape.* 11-21.

Bunting, E. S. (1986). Oilseed rape in perspective. In: *Oilseed rape.*

Burgin, L. E., Gloster, J., Sanders, C., Mellor, P. S., Gubbins, S., & Carpenter, S. (2013). Investigating incursions of bluetongue virus using a model of long-distance *Culicoides* biting midge dispersal. *Transboundary and emerging diseases*, **60**, 263-272.

Bursell, E. (1974). Environmental aspects–temperature. In: *The Physiology of Insecta* (pp. 1-41). Academic Press.

Butenandt, V. A. (1959). Uber den sexsual-lockstoff des seidenspinners Bombyx mori. *Reindarstellung und konstitution. Z. Naturforschg, b*, **14**, 283.

Butler, C. (1609). *The Feminine Monarchie. On a Treatise concerning bees, and the due ordering of them.*

Byrne, D. N. (1999). Migration and dispersal by the sweet potato whitefly, *Bemisia tabaci*. *Agricultural and forest meteorology*, **97**, 309-316.

Cai, Y. W., Cheng, X. Y., Xu, R. M., Duan, D. H., & Kirkendall, L. R. (2008). Genetic diversity and biogeography of red turpentine beetle Dendroctonus valens in its native and invasive regions. *Insect Science*, **15**, 291-301.

Callahan, J. T. (1984). Long-term ecological research. *BioScience*, **34**, 363-367.

Cammell, M. E., & Knight, J. D. (1992). Effects of climatic change on the population dynamics of crop pests. In: *Advances in Ecological Research* (Vol. 22, pp. 117-162). Academic Press.

Carpenter, S. R., Cole, J. J., Pace, M. L., Batt, R., Brock, W. A., Cline, T., Coloso, J., Hodgson, J. R., Kitchell, J. F., Seekell, D. A. & Weidel, B. (2011). Early warnings of regime shifts: a whole-ecosystem experiment. *Science*, **332**, 1079-1082.

Carreck, N. L. (2017). A beekeeper's perspective on the neonicotinoid ban. *Pest Management Science*, **73**, 1295-1298.

Carreck, N. L., & Ratnieks, F. L. (2014). The dose makes the poison: have “field realistic” rates of exposure of bees to neonicotinoid insecticides been overestimated in laboratory studies? *Journal of Apicultural Research*, **53**, 607-614.

Carrière, Y., Ellsworth, P. C., Dutilleul, P., Ellers-Kirk, C., Barkley, V., & Antilla, L. (2006). A GIS-based approach for areawide pest management: the scales of *Lygus hesperus* movements to cotton from alfalfa, weeds, and cotton. *Entomologia Experimentalis et Applicata*, **118**, 203-210.

Cartwright, B., Edelson, J. V., & Chambers, C. (1987). Composite action thresholds for the control of lepidopterous pests on fresh-market cabbage in the lower Rio Grande Valley of Texas. *Journal of Economic Entomology*, **80**, 175-181.

Cassel-Lundhagen, A., Öberg, S., Högfeldt, C., & Ekbom, B. (2009). Species-specific primers for predation studies of the pollen beetle, *Meligethes aeneus* (Coleoptera, Nitidulidae). *Molecular Ecology Resources*, **9**, 1132-1134.

Catchen, J., Amores, A., Hohenlohe, P., Cresko, W. & Postlethwait, J. (2011) Stacks: building and genotyping loci de novo from short-read sequences. *G3: Genes, Genomes, Genetics*, **1**, 171-182.

Catchen, J., Hohenlohe, P. A., Bassham, S., Amores, A., & Cresko, W. A. (2013). Stacks: an analysis tool set for population genomics. *Molecular Ecology*, **22**, 3124-3140.

Cato, A. J., Bateman, N. R., Lorenz, G. M., Hardke, J. T., Black, J. L., Thrash, B. C., Johnson, D. L., Gore, J., Studebaker, G., Fan, S. X. & Gaillard, P. R. (2019). Influence of sweep length on rice stink bug (Hemiptera: Pentatomidae) capture and reliability of population density estimates. *Journal of Economic Entomology*, **112**, 2713-2718.

Chakraborty, R. (1987). Biochemical heterozygosity and phenotypic variability of polygenic traits. *Heredity*, **59**, 19-28.

Chakraborty, R., & Nei, M. (1977). Bottleneck effects on average heterozygosity and genetic distance with the stepwise mutation model. *Evolution*, 347-356.

Champion, G. T., May, M. J., Bennett, S., Brooks, D. R., Clark, S. J., Daniels, R. E., Firbank, L. G., Haughton, A. J., Hawes, C. & Heard, M. S. (2003). Crop management and agronomic context of the Farm Scale Evaluations of genetically modified herbicide–tolerant crops. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, **358**, 1801-1818.

Chancellor, T. C. B., Cook, A. G., Heong, K. L. & S. Villareal. (1997). The flight activity and infectivity of the major leafhopper vectors (Hemiptera: Cicadellidae) of rice tungro viruses in an irrigated rice area in the Philippines. *Bulletin of Entomological Research,* **87**, 247–258.

Chapman, J. W., Bell, J. R., Burgin, L. E., Reynolds, D. R., Pettersson, L. B., Hill, J. K., Bonsall, M. B. & Thomas, J. A. (2012). Seasonal migration to high latitudes results in major reproductive benefits in an insect. *Proceedings of the National Academy of Sciences*, **109**, 14924-14929.

Chapman, J. W., Drake, V. A., & Reynolds, D. R. (2011). Recent insights from radar studies of insect flight. *Annual Review of Entomology*, **56**, 337-356.

Chapman, J. W., Reynolds, D. R., & Smith, A. D. (2003). Vertical-looking radar: a new tool for monitoring high-altitude insect migration. *Bioscience*, **53**, 503-511.

Chapman, J., Reynolds, D., & Smith, A. (2004a). Migratory and foraging movements in beneficial insects: a review of radar monitoring and tracking methods. *International Journal of Pest Management*, **50**, 225-232.

Chapman, J. W., Reynolds, D. R., Smith, A. D., Smith, E. T., & Woiwod, I. P. (2004b). An aerial netting study of insects migrating at high altitude over England. *Bulletin of Entomological Research*, **94**, 123-136.

Chapman, J. W., Reynolds, D. R., & Wilson, K. (2015). Long-range seasonal migration in insects: mechanisms, evolutionary drivers and ecological consequences. *Ecology Letters*, **18**, 287-302.

Charlesworth, D. (2003). Effects of inbreeding on the genetic diversity of populations. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, **358**, 1051-1070.

Charpentier, R. (1985). Host plant selection by the pollen beetle *Meligethes aeneus*. *Entomologia Experimentalis et Applicata*, **38**, 277-285.

Chen, S., & Andreasson, E. (2001). Update on glucosinolate metabolism and transport. *Plant Physiology and Biochemistry*, **39**, 743-758.

Chilson, P. B., Frick, W. F., Kelly, J. F., Howard, K. W., Larkin, R. P., Diehl, R. H., Westbrook, J. K., Kelly, T. A. & Kunz, T. H. (2012). Partly cloudy with a chance of migration: weather, radars, and aeroecology. *Bulletin of the American Meteorological Society*, **93**, 669-686.

Chilson, P. B., Stepanian, P. M., & Kelly, J. F. (2017). Radar aeroecology. In: *Aeroecology* (pp. 277-309). Springer, Cham.

Choi, W. I., Ryoo, M. I., Chung, Y. J., & Park, Y. S. (2011). Geographical variation in the population dynamics of *Thecodiplosis japonensis*: causes and effects on spatial synchrony. *Population Ecology*, **53**, 429-439.

Chollet, F. (2015). *Keras*. https://github.com/fchollet/keras.

Christen, O., & Sieling, K. (1995). Effect of Different Preceding Crops and Crop Rotations on Yield of Winter Oil-seed Rape (*Brassica napus* L.). *Journal of Agronomy and Crop Science*, **174**, 265-271.

Clobert, J., Wolff, J. O., Nichols, J. D., Danchin, E. & Dhondt, A. A. (2001). Introduction. In: *Dispersal* (pp xvii-xxi) Oxford University Press.

Clutton-Brock, T., & Sheldon, B. C. (2010). Individuals and populations: the role of long-term, individual-based studies of animals in ecology and evolutionary biology. *Trends in Ecology & Evolution*, **25**, 562-573.

Cockerham, C. C. (1973). Analyses of gene frequencies. *Genetics*, **74**, 679-700.

Cocu, N., Conrad, K., Harrington, R. & Rounsevell, M. D. A. (2005a) Analysis of spatial patterns at a geographical scale over north-western Europe from point-referenced aphid count data. *Bulletin of Entomological Research,* **95***,* 47-56.

Cocu, N., Harrington, R., Hullé, M. & Rounsevell, M. D. A. (2005b) Spatial autocorrelation as a tool for identifying the geographical patterns of aphid annual abundance. *Agricultural and Forest Entomology* **7**, 31-43.

Cocu, N., Harrington, R., Rounsevell, M. D. A., Worner, S. P., Hulle, M., & Examine Project Participants. (2005c). Geographical location, climate and land use influences on the phenology and numbers of the aphid, *Myzus persicae*, in Europe. *Journal of Biogeography*, **32**, 615-632.

Collier, R. H., Finch, S., Emmett, B. J., & Blood-Smyth, J. (1990). Monitoring populations of the carrot fly, *Psila rosae*, with sticky and water traps. *Annals of Applied Biology*, **116**, 441-446.

Collier, R. H., Finch, S., Endersby, N. M., & Ridland, P. M. (2001). Forecasting attacks by pest insects of cruciferous crops. In: *The Management of Diamondback Moth and Other Crucifer Pests. Proceedings of the Fourth International Workshop on Management of the Diamondback Moth and Other Crucifer Pests* (pp. 163-168). Department of Natural Resources and Environment. Melbourne, Australia.

Collins, L. (2017). *Project Report No. 572: National Survey of Cabbage Stem Flea Beetle Larvae in Winter Oilseed Rape Plants in Autumn 2016 and Spring 2017* AHDB.

Colpitts, B. G., & Boiteau, G. (2004). Harmonic radar transceiver design: Miniature tags for insect tracking. *IEEE Transactions on Antennas and Propagation*, **52**, 2825-2832.

Conrad, K. F., Warren, M. S., Fox, R., Parsons, M. S., & Woiwod, I. P. (2006). Rapid declines of common, widespread British moths provide evidence of an insect biodiversity crisis. *Biological Conservation*, **132**, 279-291.

Cook, S. (2000). *The use of pollen cues in resource location by a pollinator and a pest* (Doctoral dissertation, University of Nottingham).

Cook, S. M., Bartlett, E., Murray, D. A., & Williams, I. H. (2002). The role of pollen odour in the attraction of pollen beetles to oilseed rape flowers. *Entomologia Experimentalis et Applicata*, **104**, 43-50.

Cook, S. M., & Denholm, I. (2008). Ecological approaches to the control of pollen beetles in oilseed rape. *EPPO Bulletin*, **38**, 110-113.

Cook, S. M., Döring, T. F., Ferguson, A. W., Martin, J. L., Skellern, M. P., Smart, L. E., Watts, N. P., Welham, S. J., Woodcock, C. & Pickett, J. A. (2013a). Development of an integrated pest management strategy for control of pollen beetles in winter oilseed rape (No. 504). *HGCA project report*.

Cook, S. M., Ferguson, A. W., Skellern, M. P., Watts, N. P., Martin, J. L., Smart, L. E., Woodcock, C. M. & Pickett, J. A. (2014) Developing an integrated pest management strategy for pollen beetles in winter oilseed rape: results from the defra SA-LINK project LK09108. *IOBC-WPRS Bulletin,* **104**, 96.

Cook, S. M., Khan, Z. R., & Pickett, J. A. (2007). The use of push-pull strategies in integrated pest management. *Annu. Rev. Entomol.*, **52**, 375-400.

Cook, S. M., Murray, D. A., Watts, N. P., & Williams, I. H. (2006a). Responses of pollen beetles (*Meligethes aeneus*) to conspecific odours. *IOBC-WPRS Bulletin*, **29**, 143.

Cook, S. M., Murray, D. A., & Williams, I. H. (2004a). Do pollen beetles need pollen? The effect of pollen on oviposition, survival, and development of a flower-feeding herbivore. *Ecological Entomology*, **29**, 164-173.

Cook, S. M., Rasmussen, H. B., Birkett, M. A., Murray, D. A., Pye, B. J., Watts, N. P., & Williams, I. H. (2007). Behavioural and chemical ecology underlying the success of turnip rape (*Brassica rapa)* trap crops in protecting oilseed rape (*Brassica napus*) from the pollen beetle (*Meligethes aeneus*). *Arthropod-Plant Interactions*, **1**, 57-67.

Cook, S. M., Skellern, M. P., Döring, T. F., & Pickett, J. A. (2013b). Red oilseed rape? The potential for manipulation of petal colour in control strategies for the pollen beetle (*Meligethes aeneus*). *Arthropod-Plant Interactions*, **7**, 249-258.

Cook, S. M., Skellern, M. P., Smith, M. J., & Williams, I. H. (2006). Responses of pollen beetles (*Meligethes aeneus*) to petal colour. *IOBC-WPRS Bulletin*, **29**, 151.

Cook, S. M., Watts, N. P., Hunter, F., Smart, L. E., & Williams, I. H. (2004b). Effects of a turnip rape trap crop on the spatial distribution of *Meligethes aeneus* and *Ceutorhynchus assimilis* in oilseed rape. *IOBC-WPRS Bulletin*, **27**, 199-206.

Cooke, B. J., & Lorenzetti, F. (2006). The dynamics of forest tent caterpillar outbreaks in Quebec, Canada. *Forest Ecology and Management*, **226**, 110-121.

Cooter, J. (1977). Swarming of *Meligethes aeneus* (F.)(Coleoptera, Nitidulidae). *Entomologists' Monthly Magazine*.

Coppel, H. C., Casida, J. E., & Dauterman, W. C. (1960). Evidence for a potent sex attractant in the introduced pine sawfly, *Diprion similis* (Hymenoptera: Diprionidae). *Annals of the Entomological Society of America*, **53**, 510-512.

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

Coulthard, E., Norrey, J., Shortall, C., & Harris, W. E. (2019). Ecological traits predict population changes in moths. *Biological Conservation*, **233**, 213-219.

Crawley, M. J. (2007). *The R Book*. Chichester. *UK: Wiley*.

Cresswell, J. E., & Thompson, H. M. (2012). Comment on "A common pesticide decreases foraging success and survival in honey bees". *Science*, **337**, 1453-1453.

Crossley, M. S., Rondon, S. I., & Schoville, S. D. (2019). Effects of contemporary agricultural land cover on Colorado potato beetle genetic differentiation in the Columbia Basin and Central Sands. *Ecology and evolution*, **9**, 9385-9394.

Croteau, E. K. (2010) Causes and Consequences of Dispersal in Plants and Animals. *Nature Education Knowledge*, **3**:12.

Cutler, D. R., Edwards Jr, T. C., Beard, K. H., Cutler, A., Hess, K. T., Gibson, J., & Lawler, J. J. (2007). Random forests for classification in ecology. *Ecology*, **88**, 2783-2792.

Cytrynowicz, M., Morgente, J. S. & De Souza, H. M. L. 1982. Visual responses of South American Fruit flies Anastrepha praterulus and Mediterranean fruit flies *Ceratitis capitata* to colored rectangles and spheres. *Environmental Entomology,* **11**, 1202–1210.

Damos, P. (2015). Modular structure of web-based decision support systems for integrated pest management. A review. *Agronomy for Sustainable Development*, **35**, 1347-1372.

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T. & McVean, G. (2011). The variant call format and VCFtools. *Bioinformatics*, **27**, 2156-2158.

Darwin, C. (1896). *The descent of man and selection in relation to sex (Vol. 1).* D. Appleton.

David, G., Giffard, B., Piou, D., & Jactel, H. (2014). Dispersal capacity of *Monochamus galloprovincialis*, the European vector of the pine wood nematode, on flight mills. *Journal of Applied Entomology*, **138**, 566-576.

Day, R. K., & Knight, J. D. (1995). Operational aspects of forecasting migrant insect pests. In: *Insect Migration: tracking resources through space and time*. Cambridge: Cambridge University, 323-334.

De'ath, G., & Fabricius, K. E. (2000). Classification and regression trees: a powerful yet simple technique for ecological data analysis. *Ecology*, **81**, 3178-3192.

Dedryver, C. A., Parisey, N., Turpeau-Ait, I., & Hulle, M. (2011). Climate and aphid flight: what we get from 30 years of grain aphid *S. avenae* catches by the French suction trap network Agraphid. In: *Les Cochenilles: ravageur principal ou secondaire. 9ème Conférence Internationale sur les Ravageurs en Agriculture, SupAgro, Montpellier, France, 25-27 Octobre 2011* (pp. 550-558). Association Française de Protection des Plantes (AFPP).

DEFRA (2014) Farming Statistics Provisional crop areas, yields and livestock populations At June 2014 – United Kingdom https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/364157/structure-jun2013prov-UK-16oct14.pdf accessed October 2015.

DEFRA (2018) Farming Statistics Final crop areas, yields, livestock populations and agricultural workforce At 1 June 2018 - United Kingdom https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/805472/structure-jun2018final-eng-31may19.pdf accessed January 2020.

DEFRA (2019a) Farming Statistics Final crop areas, yields, livestock populations and agricultural workforce At June 2019 - United Kingdom https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/853450/structure-jun2019final-uk-19dec19.pdf accessed January 2020.

DEFRA (2019b) *Butterflies in England: species of the wider countryside on farmland and in woodland, 1990 to 2018* https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/815771/Butterflies_in_England_1990_2018.pdf accessed January 2020.

DEFRA (2020) Early Bird Survey 2020 https://ahdb.org.uk/cereals-oilseeds/early-bird-survey accessed January 2020de Freitas Bueno, R. C. O., de Freitas Bueno, A., Moscardi, F., Postali Parra, J. R., & Hoffmann-Campo, C. B. (2011). Lepidopteran larva consumption of soybean foliage: basis for developing multiple-species economic thresholds for pest management decisions. *Pest Management Science*, **67**, 170-174.

Defriez, E. J., Sheppard, L. W., Reid, P. C., & Reuman, D. C. (2016). Climate change-related regime shifts have altered spatial synchrony of plankton dynamics in the North Sea. *Global Change Biology*, **22**, 2069-2080.

de la Pasture, L. (2016). OSR planning – Beetle larvae may force rethink. *Crop Production Magazine* (https://www.cpm-magazine.co.uk/2016/05/20/osr-planning-beetle-larvae-may-force-rethink/ accessed December 2020).

Delignette-Muller, M. L., & Dutang, C. (2015). Fitdistrplus: An R Package for fitting distributions. *Journal of Statistical Software*, **64**, 1982– 34.

Denholm, I., Sawicki, R. M., & Farnham, A. W. (1985). Factors affecting resistance to insecticides in house-flies, *Musca domestica* L.(Diptera: Muscidae). IV. The population biology of flies on animal farms in south-eastern England and its implications for the management of resistance. *Bulletin of Entomological Research*, **75**, 291-304.

Dennis, B., Desharnais, R. A., Cushing, J. M., Henson, S. M. & Constantino, R. F. (2001) Estimating chaos and complex dynamics in an insect population *Ecological Monographs,* **71**, 277-303.

Denny, M. W., Hunt, L. J., Miller, L. P., & Harley, C. D. (2009). On the prediction of extreme ecological events. *Ecological Monographs*, **79**, 397-421.

De Oliveira Aparecido, L. E., de Souza Rolim, G., De, J. R. D. S. C., Costa, C. T. S., & de Souza, P. S. (2020). Machine learning algorithms for forecasting the incidence of *Coffea arabica* pests and diseases. *International Journal of Biometeorology*, 1-18.

de Roos, A. M., Mccauley, E., & Wilson, W. G. (1991). Mobility versus density-limited predator-prey dynamics on different spatial scales. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, **246**, 117-122.

de Roos, A. M., McCauley, E., & Wilson, W. G. (1998). Pattern formation and the spatial scale of interaction between predators and their prey. *Theoretical Population Biology*, **53**, 108-130.

Desender, K., Backeljau, T., Delahaye, K., & De Meester, L. (1998). Age and size of European saltmarshes and the population genetic consequences for ground beetles. *Oecologia*, **114**, 503-513.

Desender, K., Small, E., Gaublomme, E., & Verdyck, P. (2005). Rural–urban gradients and the population genetic structure of woodland ground beetles. *Conservation Genetics*, **6**, 51-62.

Desjardins-Proulx, P., Laigle, I., Poisot, T., & Gravel, D. (2017). Ecological interactions and the Netflix problem. *PeerJ*, **5**, e3644.

Dewar, A. M. (2017). The adverse impact of the neonicotinoid seed treatment ban on crop protection in oilseed rape in the United Kingdom. *Pest Management Science*, **73**, 1305-1309.

Diepenbrock, W. (2000). Yield analysis of winter oilseed rape (*Brassica napus* L.): a review. *Field Crops Research*, **67**, 35-49.

Dingle, H. (2014). *Migration: the biology of life on the move*. Oxford University Press, USA.

Dinsdale, E. A., Edwards, R. A., Bailey, B., Tuba, I., Akhter, S., McNair, K., Schmieder, R., Apkarian, N., Creek, M., Guan, E. & Hernandez, M. (2013). Multivariate analysis of functional metagenomes. *Frontiers in Genetics*, **4**, 41.

Dirzo, R., Young, H. S., Galetti, M., Ceballos, G., Isaac, N. J., & Collen, B. (2014). Defaunation in the Anthropocene. *Science*, **345**, 401-406.

Dlabola, J., & Taimr, L. (1965). Some results obtained with the application of the tracer method in insect migration and dispersion studies. *Acta Entomologica Bohemoslovaca*, **62**, 413-420.

Doak, D. F., Estes, J. A., Halpern, B. S., Jacob, U., Lindberg, D. R., Lovvorn, J., Monson, D. H., Tinker, M. T., Williams, T. M., Wootton, J. T. & Novak, M. (2008). Understanding and predicting ecological dynamics: are major surprises inevitable. *Ecology*, **89**, 952-961.

Doitsidis, L., Fouskitakis, G. N., Varikou, K. N., Rigakis, I. I., Chatzichristofis, S. A., Papafilippaki, A. K., & Birouraki, A. E. (2017). Remote monitoring of the *Bactrocera oleae* (Gmelin)(Diptera: Tephritidae) population using an automated McPhail trap. *Computers and Electronics in Agriculture*, **137**, 69-78.

Döring, T. F., Skellern, M., Watts, N., & Cook, S. M. (2012). Colour choice behaviour in the pollen beetle *Meligethes aeneus* (Coleoptera: Nitidulidae). *Physiological Entomology*, **37**, 360-378.

Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., Marquéz, J. R. G., Gruber, B., Lafourcade, B., Leitao, P. J. & Münkemüller, T. (2013). Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography*, **36**, 27-46.

Dormann, C. F., McPherson, J. M., Araújo, M. B., Bivand, R., Bolliger, J., Carl, G., Davies, R. G., Hirzel, A., Jetz, W., Daniel Kissling, W. & Kühn, I. (2007). Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. *Ecography*, **30**, 609–628.

Dosdall, L. M., & Mason, P. G. (2010). Key pests and parasitoids of oilseed rape or canola in North America and the importance of parasitoids in integrated management. In: *Biocontrol-based integrated management of oilseed rape pests* (pp. 167-213). Springer Netherlands.

Doty, R. L. (2010). *The great pheromone myth*. JHU press.

Downey, R. K. (1987). Genetic manipulation of oilseed quality. In: *The metabolism, structure, and function of plant lipids* (pp. 669-676). Springer, Boston, MA.

Downey, R. K., & Craig, B. M. (1964). Genetic control of fatty acid biosynthesis in rapeseed (*Brassica napus* L.). *Journal of the American Oil Chemists' Society*, **41**, 475-478.

Drake, V. A., & Reynolds, D. R. (2012). *Radar entomology: observing insect flight and migration*. Cabi.

Drake, V. A., & Wang, H. (2013). Recognition and characterization of migratory movements of Australian plague locusts, *Chortoicetes terminifera*, with an insect monitoring radar. *Journal of Applied Remote Sensing*, **7**, 075095.

Du, Y., Poppy, G. M., Powell, W., Pickett, J. A., Wadhams, L. J., & Woodcock, C. M. (1998). Identification of semiochemicals released during aphid feeding that attract parasitoid *Aphidius ervi*. *Journal of Chemical Ecology*, **24**, 1355-1368.

Dumont, M. J., & Narine, S. S. (2008). Characterization of soapstock and deodorizer distillates of vegetable oils using gas chromatography. *Lipid Technology*, **20**, 136-138.

Dunning Jr, J. B., Stewart, D. J., Danielson, B. J., Noon, B. R., Root, T. L., Lamberson, R. H., & Stevens, E. E. (1995). Spatially explicit population models: current forms and future uses. *Ecological Applications*, **5**, 3-11.

Eagles, D., Melville, L., Weir, R., Davis, S., Bellis, G., Zalucki, M. P., Walker, P. J. & Durr, P. A. (2014). Long-distance aerial dispersal modelling of *Culicoides* biting midges: case studies of incursions into Australia. *BMC Veterinary Research*, **10**, 135.

Edwards, O. R., Franzmann, B., Thackray, D., & Micic, S. (2008). Insecticide resistance and implications for future aphid management in Australian grains and pastures: a review. *Australian Journal of Experimental Agriculture*, **48**, 1523-1530.

Efron, B., & Tibshirani, R. J. (1993). *An introduction to the bootstrap*. Chapman and Hall, London, UK.

Eickermann, M., Ulber, B., Hoffmann, L., & Junk, J. (2014). Improving phenological forecasting models for rape stem weevil, *Ceutorhynchus napi* Gyll., based on long-term multisite datasets. *Journal of Applied Entomology*, **138**, 754-762.

Ekbom, B. (1998). Clutch size and larval performance of pollen beetles on different host plants. *Oikos*, 56-64.

Ekbom, B. (2010). Pests and Their Enemies in Spring Oilseed Rape in Europe and Challenges to Integrated Pest management. In: *Biocontrol-based integrated management of oilseed rape pests* (pp. 151-165). Springer Netherlands.

Ekbom, B. & Borg, A. (1996). Pollen beetle (Meligethes aeneus) oviposition and feeding preference on different host plant species. *Entomologia Experimentalis et Applicata*, **78**, 291-299.

Ekbom, B. & Ferdinand, V. (2003). Field oviposition rates and egg load dynamics of pollen beetles (*Meligethes aeneus* Fab.)(Coleoptera: Nitidulidae). *Agricultural and Forest Entomology*, **5**, 247-252.

Ekbom, B. & Popov, S. Y. (2004). Host plant affects pollen beetle (*Meligethes aeneus*) egg size. *Physiological Entomology*, **29**, 118-122.

Elkington, J. S., & Carde, R. T. (1984). Effect of wild and laboratory-reared female gypsy moth (Lepidoptera: Lymantriidae) on the capture of males in pheromone-baited traps. *Environmental Entomol*ogy **13**, 1377-1385.

Elliott, D. P. & Powell, J. M. (1966) A cage for collecting insects from tree stems and branches. *Canadian Entomologist*, **98**, 1112-1113.

Endersby, N. M., & Morgan, W. C. (1991). Alternatives to synthetic chemical insecticides for use in crucifer crops. *Biological Agriculture & Horticulture*, **8**, 33-52.

Engen, S., & Sæther, B. E. (2005). Generalizations of the Moran effect explaining spatial synchrony in population fluctuations. *The American Naturalist*, **166**, 603-612.

Englund, R. A. (2001). Long-term monitoring of one of the most restricted insect populations in the United States, *Megalagrion xanthomelas* (Selys-Longchamps), at Tripler Army Medical Center, Oahu, Hawaii (Zygoptera: Coenagrionidae). *Odonatologica*, **30**, 255-263.

Estay, S. A., Lima, M., & Harrington, R. (2009). Climate mediated exogenous forcing and synchrony in populations of the oak aphid in the UK. *Oikos*, **118**, 175-182.

Evans, K. A., & Allen-Williams, L. J. (1994). Laboratory and field response of the pollen beetle, *Meligethes aeneus*, to the odour of oilseed rape. *Physiological Entomology*, **19**, 285-290.

Fabre, F. & Dedryver, C-A. (2017). Decision Support Systems. In: *Aphids as crop pests* (pp. 382-397). Cabi.

Fabre, F., Dedryver, C-A., Plantegenest, M., Hullé, M. & Rivot, E. (2010) Hierarchical Bayesian Modelling of plant colonisation by winged aphids: Inferring dispersal processes by linking aerial and field count data. *Ecological Modelling* **221**, 1770-1778.

Fabre, J. H. (1879). *Souvenirs Entomologiques* Delagrave.

FAO (2021) *Pests of economic importance in Ukraine: Integrated pest management manual.* Budapest.

Farazmand, M., & Sapsis, T. P. (2019). Extreme events: Mechanisms and prediction. *Applied Mechanics Reviews*, **71**, 050801.

Farrell, G. (2000). Dispersal, phenology and predicted abundance of the larger grain borer in different environments. *African Crop Science Journal*, **8**, 337-343.

Fassotte, C., Delécolle, J. C., Cors, R., Defrance, T., De Deken, R., Haubruge, E., & Losson, B. (2008). *Culicoides* trapping with Rothamsted suction traps before and during the bluetongue epidemic of 2006 in Belgium. *Preventive Veterinary Medicine*, **87**, 74-83.

Fedorenko, V. P. (2006). The most important sugar beet pests in Ukraine and integral measures for their control. *Zbornik Matice srpske za prirodne nauke*, **110**, 21-38.

Feener Jr., D. H., Jacobs, L. F., & Schmidt, J. O. (1996). Specialized parasitoid attracted to a pheromone of ants. *Animal Behaviour*, **51**, 61-66. Appendix.

Ferguson, A. W., Barari, H., Warner, D. J., Campbell, J. M., Smith, E. T., Watts, N. P. & Williams, I. H. (2006) Distributions and interactions of the stem miners *Psylliodes chrysocephala* and *Ceutorhynchus pallidactylus* and their parasitoids in a crop of winter oilseed rape (*Brassica napus*). *Entomologia Experimentalis et Applicata*, **119**, 81-92.

Ferguson, A. W., Campbell, J. M., Warner, D. J., Watts, N. P., Schmidt, J. E. U., & Williams, I. H. (2003a). Spatio-temporal distributions of *Meligethes aeneus* and its parasitoids in an oilseed rape crop and their significance for crop protection. *Proc 11th Int Rapeseed Cong, 6–10 July 2003, Copenhagen, Denmark*, 1057-1059.

Ferguson, A. W., Campbell, J. M., Warner, D. J., Watts, N. P., Schmidt, J. E. U. & Williams, I. H. (2004) Phenology and spatial distributions of Dasineura brassicae and its parasitoids in a crop of winter oilseed rape: implications for integrated pest management. *Integrated Protection in Oilseed Crops IOBC-WPRS Bulletin*, **27**, 243-251.

Ferguson, A. W., Holdgate, R., Mason, N. S., Clark, S. J., & Williams, I. H. (2013). Phenologies and diel periodicities of within-crop flight by pests and parasitoids in winter oilseed rape in the UK. *IOBC-WPRS Bulletin*, **92**, 45-54.

Ferguson, A. W., Klukowski, Z., Walczak, B., Clark, S. J., Mugglestone, M. A., Perry, J. N., & Williams, I. H. (2003b). Spatial distribution of pest insects in oilseed rape: implications for integrated pest management. *Agriculture, Ecosystems & Environment*, **95**, 509-521.

Ferguson, A. W., Nevard, L. M., Clark, S. J., & Cook, S. M. (2015). Temperature-activity relationships in *Meligethes aeneus*: implications for pest management. *Pest Management Science*, **71**, 459-466.

Ferguson, A. W., Skellern, M. P., Johnen, A., von Richthofen, J. S., Watts, N. P., Bardsley, E., Murray, D. A. & Cook, S. M. (2016). The potential of decision support systems to improve risk assessment for pollen beetle management in winter oilseed rape. *Pest Management Science*, **72**, 609-617.

Ferguson, A. W., Williams, I. H., Castle, L. M., & Skellern, M. (2010). Key parasitoids of the pests of oilseed rape in Europe: A guide to their identification. In: *Biocontrol-Based Integrated Management of Oilseed Rape Pests* (pp. 77-114). Springer, Dordrecht.

Fernández García, A., Griffiths, G. J. K. & Thomas, C. F. G. (2000). Density, distribution and dispersal of the carabid beetle *Nebria brevicollis* in two adjacent cereal fields. *Annals of Applied Biology,* **137**, 1–9.

Finch, S., & Collier, R. H. (1989). Effects of the angle of inclination of traps on the numbers of large Diptera caught on sticky boards in certain vegetable crops. *Entomologia Experimentalis et Applicata*, **52**, 23-27.

Finch, S., Collier, R. H., & Elliott, M. S. (1990). Seasonal variations in the timing of attacks of bronzed blossom beetles (*Meligethes aeneus/Meligethes viridescens)* on horticultural brassicas. In: *Brighton Crop Protection Conference, Pests and Diseases-1990. Vol. 1.* (pp. 349-354). British Crop Protection Council.

Finch, S., Collier, R. H., & Phelps, K. (1996). A review of work done to forecast pest insect attacks in UK horticultural crops. *Crop Protection*, **15**, 353-357.

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.

Fletcher, R., & Fortin, M. (2018). *Spatial ecology and conservation modeling*. Springer International Publishing.

Fournier, F., Pelletier, D., Vigneault, C., Goyette, B., & Boivin, G. (2005). Effect of barometric pressure on flight initiation by *Trichogramma pretiosum* and *Trichogramma evanescens (*Hymenoptera: Trichogrammatidae). *Environmental Entomology*, **34**, 1534-1540.

Fowler, W. W. (1885). The Nitidulae of Great Britain. *Entomologist's Monthly Magazine,* 21, 54-58, 92-98, 142-147, 213-219, 260-267.

Fox, R. (2013). The decline of moths in Great Britain: a review of possible causes. *Insect Conservation and Diversity*, **6**, 5-19.

Frank, T., Drapela, T., Moser, D., & Zaller, J. G. (2010). Insect pests and spiders in oilseed rape and their response to site and landscape factors. In: *Biocontrol-Based Integrated Management of Oilseed Rape Pests* (pp. 285-304). Springer, Dordrecht.

Franklin, J. F. (1989). Importance and justification of long-term studies in ecology. In: *Long-term studies in ecology* (pp. 3-19). Springer, New York, NY.

Frearson, D. (2006). *The movement and host-selection of three inflorescence insect pests of oilseed rape and their parasitoids in relation to host-plant characteristics and environmental conditions* PhD thesis, University of Reading.

Frearson, D. J., Ferguson, A. W., Campbell, J. M., & Williams, I. H. (2005). The spatial dynamics of pollen beetles in relation to inflorescence growth stage of oilseed rape: implications for trap crop strategies. *Entomologia Experimentalis et Applicata*, **116**, 21-29.

Frearson, D. J. T., Ferguson, A. W., Laugier, N., & Williams, I. H. (2006). The influence of petals on numbers of pests and their parasitoids in oilseed rape. In: *Proc Symp Integrated Pest Management in Oilseed Rape, 3–5 April 2006, Göttingen, Germany*.

Free, J. B., & Williams, I. H. (1978). The responses of the pollen beetle, *Meligethes aeneus*, and the seed weevil, *Ceuthorhynchus assimilis*, to oil-seed rape, Brassica napus, and other plants. *Journal of Applied Ecology*, 761-774.

Friederichs, K. (1921). Untersuchungen über den Rapsglanzkafers in Mecklenburg. *Z. angew. Ent. Berl*., **7**, 1-36.

Friedt, W. & Snowdon, R. (2009). Oilseed rape. In: *Oil crops* (pp. 91-126). Springer, New York, NY.

Friedt, W., Tu, J., & Fu, T. (2018). Academic and economic importance of *Brassica napus* rapeseed. In: *The Brassica napus genome* (pp. 1-20). Springer, Cham.

Fritzsche, R. (1957). Zur Biologie und Ökologie der Rapsschädlinge aus der Gattung *Meligethes*. *Zeitschrift für angewandte Entomologie*, **40**, 222-280.

Fussell, G. E. (1955). History of cole (*Brassica sp.*). *Nature,* **176**, 48-51.

Gangwere, S. K., Chavin, W., & Evans, F. C. (1964). Methods of marking insects, with especial reference to Orthoptera (Sens. Lat.). *Annals of the Entomological Society of America*, **57**, 662-669.

Gehringer, A., Snowdon, R., Spiller, T., Basunanda, P., & Friedt, W. (2007). New oilseed rape (*Brassica napus*) hybrids with high levels of heterosis for seed yield under nutrient-poor conditions. *Breeding Science*, **57**, 315-320.

Gevrey, M., Dimopoulos, I., & Lek, S. (2006). Two-way interaction of input variables in the sensitivity analysis of neural network models. *Ecological Modelling*, **195**, 43-50.

Giamoustaris, A., & Mithen, R. (1996). The effect of flower colour and glucosinolates on the interaction between oilseed rape and pollen beetles. In: *Proceedings of the 9th International Symposium on Insect-Plant Relationships* (pp. 206-208). Springer Netherlands.

Gibbs, R. B. (2000). Long-term treatment with estrogen and progesterone enhances acquisition of a spatial memory task by ovariectomized aged rats. *Neurobiology of Aging*, **21**, 107-116.

Gill, R. J., Ramos-Rodriguez, O., & Raine, N. E. (2012). Combined pesticide exposure severely affects individual-and colony-level traits in bees. *Nature*, **491**, 105-108.

Giska, I., Babik, W., van Gestel, C. A., van Straalen, N. M., & Laskowski, R. (2015). Genome-wide genetic diversity of rove beetle populations along a metal pollution gradient. *Ecotoxicology and Environmental Safety*, **119**, 98-105.

Gitzen, R. A., Millspaugh, J. J., Cooper, A. B., & Licht, D. S. (Eds.). (2012). *Design and analysis of long-term ecological monitoring studies*. Cambridge University Press.

Glen, D. M. (1976) An emergence trap for bark-dwelling insects, its efficiency and effects on temperature. *Ecological Entomology*, **1**, 91-94.

Glick, P. A. (1939). The distribution of insects, spiders, and mites in the air *U.S.D.A. Technical Bulletin no. 673*.

Godfrey, J. D., & Bryant, D. M. (2003). Effects of radio transmitters: Review of recent radio-tracking studies. *Science for Conservation*, **214**, 83-95.

González-Andújar, J. L. (2020). Introduction to Decision Support Systems. In: *Decision Support Systems for Weed Management* (pp. 25-38). Springer, Cham.

Goodwin, P., & Wright, G. (2010). The limits of forecasting methods in anticipating rare events. *Technological Forecasting and Social Change*, **77**, 355-368.

Græsbøll, K., Enøe, C., Bødker, R., & Christiansen, L. E. (2014). Optimal vaccination strategies against vector-borne diseases. *Spatial and Spatio-temporal Epidemiology*, **11**, 153-162.

Graham, H. M., Wolfenbarger, D. A., Nosky, J. R., Hernandez Jr, N. S., Llanes, J. R., & Tamayo, J. A. (1978). Use of rubidium to label corn earworm and fall armyworm for dispersal studies. *Environmental Entomology*, ***7***, 435-438.

Grapputo, A. (2006). Development and characterization of microsatellite markers in the Colorado potato beetle, *Leptinotarsa decemlineata*. *Molecular Ecology Notes*, **6**, 1177-1179.

Gressitt, J. L., Coatsworth, J., & Yoshimoto, C. M. (1962). Airborne insects trapped on 'monsoon expedition'. *Pacific Insects*, **4**, 319-323.

Griffiths, G. J., Alexander, C. J., Birt, A., Holland, J. M., Kennedy, P. J., Perry, J. N., Preston, R. & Winder, L. (2005). A method for rapidly mass laser-marking individually coded ground beetles (Coleoptera: Carabidae) in the field. *Ecological Entomology*, **30**, 391-396.

Guikema, S. D., & Quiring, S. M. (2012). Hybrid data mining-regression for infrastructure risk assessment based on zero-inflated data. *Reliability Engineering & System Safety*, **99**, 178-182.

Gurney, W. S. C., Veitch, A. R., Cruickshank, I., & McGeachin, G. (1998). Circles and spirals: population persistence in a spatially explicit predator–prey model. *Ecology*, **79**, 2516-2530.

Habekotté, B. (1997a). A model of the phenological development of winter oilseed rape (*Brassica napus* L.). *Field Crops Research*, **54**, 127-136.

Habekotte, B. (1997b). Evaluation of seed yield determining factors of winter oilseed rape (*Brassica napus* L.) by means of crop growth modelling. *Field Crops Research*, **54**, 137-151.

Habekotté, B. (1997c). Identification of strong and weak yield determining components of winter oilseed rape compared with winter wheat. *European Journal of Agronomy*, **7**, 315-321.

Hadley, M. (1969) The adult biology of the crane-fly *Molophilus ater* Meigen. *Journal of Animal Ecology*, **38**, 765-790.

Hagler, J. R., & Jackson, C. G. (2001). Methods for marking insects: current techniques and future prospects. Annual Review of Entomology, **46**, 511-543.

Hameed, A., Sohail, M., Ahmad, S. S., & Karar, H. (2015). Forecasting and Modelling of *Helicoverpa armigera* (Hub.) in Relation to Weather Parameter in Multan Punjab Pakistan. *Pakistan Journal of Zoology*, **47**.

Hansen, B. B., Grøtan, V., Herfindal, I., & Lee, A. M. (2020). The Moran effect revisited: spatial population synchrony under global warming. *Ecography*, **43**, 1591-1602.

Hansen, L. M. (2003). Insecticide-resistant pollen beetles (*Meligethes aeneus* F) found in Danish oilseed rape (*Brassica napus* L) fields. *Pest Management Science*, **59**, 1057-1059.

Hansen, L. M. (2004). Economic damage threshold model for pollen beetles (*Meligethes aeneus* F.) in spring oilseed rape (*Brassica napus* L.) crops. *Crop Protection*, **23**, 43-46.

Hansen, L. M. (2008). Occurrence of insecticide resistant pollen beetles (*Meligethes aeneus* F.) in Danish oilseed rape (*Brassica napus* L.) crops. *EPPO bulletin*, **38**, 95-98.

Hanski, I., & Woiwod, I. P. (1993). Spatial synchrony in the dynamics of moth and aphid populations. *Journal of Animal Ecology*, **62**, 656-668.

Hao, P., Zhan, Y., Wang, L., Niu, Z., & Shakir, M. (2015). Feature selection of time series MODIS data for early crop classification using random forest: A case study in Kansas, USA. *Remote Sensing*, **7**, 5347-5369.

Hardie, J., & Minks, A. K. (1999). *Pheromones of non-lepidopteran insects associated with agricultural plants*. CABI publishing.

Hardy, A. C., & Cheng, L. (1986). Studies in the distribution of insects by aerial currents. III. Insect drift over the sea. *Ecological Entomology*, **11**, 283-290.

Hardy, A. C., & Milne, P. S. (1937). Insect drift over the North Sea. *Nature*, **139**, 510-511.

Hardy, A. C., & Milne, P. S. (1938). Studies in the distribution of insects by aerial currents. *The Journal of Animal Ecology*, **7**, 199-229.

Harenberg, A. (1997). *Auswirkungen extensiv geführter Anbausysteme in verschiedenen Fruchtfolgen (Raps-, Zuckerrübenfruchtfolge) und einer selbsbegrünenden Dauerbrache auf spinnen (Arachnida: Araneae)* (Doctoral dissertation, PhD thesis, Technical University of Braunschweig, Braunschweig, Germany).

Harrell Jr, F. E. (2015). *Regression modelling strategies: with applications to linear models, logistic and ordinal regression, and survival analysis*. Springer.

Harrington, R. (2014) The Rothamsted Insect Survey strikes gold. *Antenna*, **38**, 159–166.

Harrington, R., Taylor, M. S., Shortall, C. R., Alderson, L., Mallott, M., & Verrier, P. J. (2012). The Rothamsted Insect Survey: old traps, new tricks. *Aspects of Applied Biology*, **117**, 157-164.

Harrington, R., & Woiwod, I. (2007). Foresight from hindsight: the Rothamsted Insect Survey. *Outlooks on Pest Management*, **18**, 9.

Harris, A. M., & DeGiorgio, M. (2017). An unbiased estimator of gene diversity with improved variance for samples containing related and inbred individuals of any ploidy. *G3: Genes, Genomes, Genetics*, **7**, 671-691.

Hassall, K. L., Dye, A., Potamitis, I., & Bell, J. R. (2021). Resolving the identification of weak-flying insects during flight: a coupling between rigorous data processing and biology. *Agricultural and Forest Entomology*. DOI: 10.1111/afe.12453

Hatch, M. H. (1957). The North American status of *Meligethes nigrescens* Steph.(Nitidulidae). *The Coleopterists' Bulletin*, 65-66.

Haynes, K. J., Walter, J. A., & Liebhold, A. M. (2019). Population spatial synchrony enhanced by periodicity and low detuning with environmental forcing. *Proceedings of the Royal Society B*, **286**, 20182828.

Heathcote, G. D. (1957). The comparison of yellow cylindrical, flat and water traps, and of Johnson suction traps, for sampling aphids. *Annals of Applied Biology*, **45**, 133-139.

Hebert, P. D., Ratnasingham, S., & De Waard, J. R. (2003). Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, **270**, S96-S99.

Hedin, J., & Ranius, T. (2002). Using radio telemetry to study dispersal of the beetle *Osmoderma eremita*, an inhabitant of tree hollows. *Computers and Electronics in Agriculture*, **35**, 171-180.

Heimbach, U., & Brandes, M. (2016). Pyrethroid resistance of insect pests in oilseed rape in Germany since 2005. *IOBC-WPRS Bulletin*, **116**, 17-22.

Heino, M., Kaitala, V., Ranta, E., & Lindström, J. (1997). Synchronous dynamics and rates of extinction in spatially structured populations. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, **264**, 481-486.

Hendrichs J., Kenmore P., Robinson A. S., Vreysen M. J. B. (2007): Area-wide integrated pest management (AW-IPM): principles, practice and prospects. In: *Area Wide Control of Insect Pests. From Research to Field Implementation*. Springer, Dordrecht: 3–33.

Hendricks, D. E. (1985). Portable electronic detector system used with inverted-cone sex pheromone traps to determine periodicity and moth captures. *Environmental Entomology*, **14**, 199-204.

Hendricks, D. E. (1990). Electronic system for detecting trapped boll weevils in the field and transferring incident information to a computer. *Southwestern Entomologist*, **15**, 39-48.

Hendrix III, W. H., Mueller, T. F., Phillips, J. R., & Davis, O. K. (1987). Pollen as an indicator of long-distance movement of *Heliothis zea* (Lepidoptera: Noctuidae). *Environmental Entomology*, **16**, 1148-1151.

Henry, M., Beguin, M., Requier, F., Rollin, O., Odoux, J. F., Aupinel, P., Aptel, J., Tchamitchian, S. & Decourtye, A. (2012). A common pesticide decreases foraging success and survival in honey bees. *Science*, **336**, 348-350.

Hervé, M. R., & Cortesero, A. M. (2016). Potential for oilseed rape resistance in pollen beetle control. *Arthropod-Plant Interactions*, **10**, 463-475.

HGCA (2013) Monitoring and control of pollen beetles in oilseed rape. *Information sheet 18/Spring 2013* HGCA Publications, Stoneleigh Park, UK.

Hicks, B. (2018). Digital technologies at the forefront of Agritechnica. *International Pest Control*, **60**, 106-108.

Hiiesaar, K., Metspalu, L., Lääniste, P., Jõgar, K., Kuusik, A., & Jõudu, J. (2003). Insect pests on winter oilseed rape studied by different catching methods. *Agronomy Research*, **1**, 17-29.

Hillocks, R. J. (2012). Farming with fewer pesticides: EU pesticide review and resulting challenges for UK agriculture. *Crop Protection*, **31**, 85-93.

Hirschi, M., Spirig, C., Weigel, A. P., Calanca, P., Samietz, J., & Rotach, M. W. (2012). Monthly weather forecasts in a pest forecasting context: downscaling, recalibration, and skill improvement. *Journal of Applied Meteorology and Climatology*, **51**, 1633-1638.

Hitchens, N. M., Brooks, H. E., & Kay, M. P. (2013). Objective limits on forecasting skill of rare events. *Weather and Forecasting*, **28**, 525-534.

Hixon, M. A. (2008). Carrying capacity. In:.*Encyclopedia of ecology*.pp. 528-530. Newnes.

Hobbs, S. E., & Wolf, W. W. (1989). An airborne radar technique for studying insect migration. *Bulletin of Entomological Research*, **79**, 693-704.

Hoebeke, E. R., & Wheeler Jr, A. G. (1996). *Meligethes viridescens* (F.)(Coleoptera: Nitidulidae) in Maine, Nova Scotia, and Prince Edward Island: diagnosis, distribution, and bionomics of a Palearctic species new to North America. *Proceedings of the Entomological Society of Washington*, **98**, 221-227.

Hokkanen, H. M. T. (1989). Biological and agrotechnical control of the rape blossom beetle *Meligethes aeneus* (Coleoptera: Nitidulidae). *Acta Entomol. Fenn*. **53**, 25–30.

Hokkanen, H. M. T. (1993). Overwintering survival and spring emergence in *Meligethes aeneus*: effects of body weight, crowding, and soil treatment with *Beauveria bassiana*. *Entomologia Experimentalis et Applicata*, **67**, 241-246.

Hokkanen, H. M. (2000). The making of a pest: recruitment of *Meligethes aeneus* onto oilseed Brassicas. *Entomologia Experimentalis et Applicata*, **95**, 141-149.

Hokkanen, H. M. (2004). Impact of predators on pollen beetle *Meligethes aeneus* on rapeseed in Finland. *IOBC-WPRS Bulletin*, **27**, 293-296.

Hokkanen, H. M., Granlund, H., Husberg, G. B. & Markkula, M. (1986). Trap crops used successfully to control Meligethes aeneus (Col., Nitidulidae), the rape blossom beetle. *Annu. Entomol. Fennica*, **52**, 115-120.

Hokkanen, H., Husberg, G. B., & Söderblom, M. (1988). Natural enemy conservation for the integrated control of the rape blossom beetle *Meligethes aeneus* F. In: *Annales Agriculturae Fenniae*, **27**, 281-194.

Holliday, J. A., Wang, T., & Aitken, S. (2012). Predicting adaptive phenotypes from multilocus genotypes in Sitka spruce (*Picea sitchensis*) using random forest. *G3: Genes, Genomes, Genetics*, **2**, 1085-1093.

Holling, C. S. (1961). Principles of insect predation. *Annual Review of Entomology*, **6**, 163-182.

Holzworth, D. P., Snow, V., Janssen, S., Athanasiadis, I. N., Donatelli, M., Hoogenboom, G., White, J. W. & Thorburn, P. (2015). Agricultural production systems modelling and software: current status and future prospects. *Environmental Modelling & Software*, **72**, 276-286.

Hopkins, R. J., & Ekbom, B. (1996). Low oviposition stimuli reduce egg production in the pollen beetle *Meligethes aeneus*. *Physiological Entomology*, **21**, 118-122.

Horikoshi, M., & Tang, Y. (2016). *Ggfortify: Data visualization tools for statistical analysis results*. See: https://CRAN.R-project.org/package=ggfortify.

Howard, S. D., & Bickford, D. P. (2014). Amphibians over the edge: silent extinction risk of Data Deficient species. *Diversity and Distributions*, **20**, 837-846.

Hughes, B. B., Beas-Luna, R., Barner, A. K., Brewitt, K., Brumbaugh, D. R., Cerny-Chipman, E. B., Close, S. L., Coblentz, K. E., De Nesnera, K. L., Drobnitch, S. T. & Figurski, J. D. (2017). Long-term studies contribute disproportionately to ecology and policy. *BioScience*, **67**, 271-281.

Hughes, J., Reay, G., & Watson, J. (2014). Insecticide use on Scottish oilseed rape crops: historical use patterns and pest control options in the absence of neonicotinoid seed treatments. In: *Proceedings Crop Protection in Northern Britain* (pp. 21-26).

Humphries, G. R., Che-Castaldo, C., Bull, P. J., Lipstein, G., Ravia, A., Carrión, B., Bolton, T., Ganguly, A. & Lynch, H. J. (2018). Predicting the future is hard and other lessons from a population time series data science competition. *Ecological Informatics*, **48**, 1-11.

Huseth, A. S., Frost, K. E., Knuteson, D. L., Wyman, J. A., & Groves, R. L. (2012). Effects of landscape composition and rotation distance on *Leptinotarsa decemlineata* (Coleoptera: Chrysomelidae) abundance in cultivated potato. *Environmental Entomology*, **41**, 1553-1564.

Hyndman, R., Koehler, A. B., Ord, J. K., & Snyder, R. D. (2008). *Forecasting with exponential smoothing: the state space approach*. Springer Science & Business Media.

Ims, R. A., & Steen, H. (1990). Geographical synchrony in microtine population cycles: a theoretical evaluation of the role of nomadic avian predators. *Oikos*, **57**, 381-387.

IRAC. (2006). IRAC susceptibility test methods series, method no:11 Pollen Beetle *Meligethes aeneus*, adults, synthetic pyrethroids.

Jackman S. (2020). *Pscl: Classes and Methods for R Developed in the Political Science Computational Laboratory*. United States Studies Centre, University of Sydney, Sydney, New South Wales, Australia. R package version 1.5.5, https://github.com/atahk/pscl/.

Jackson, J. K., McElravy, E. R. I. C. P., & Resh, V. N. H. (1999). Long-term movements of self-marked caddisfly larvae (Trichoptera: Sericostomatidae) in a California coastal mountain stream. *Freshwater Biology*, **42**, 525-536.

Jacobson, M., & Beroza, M. (1964). Insect attractants. *Scientific American*, **211**, 20-27.

Jacobson, B., & Peres-Neto, P. R. (2010). Quantifying and disentangling dispersal in metacommunities: how close have we come? How far is there to go? *Landscape Ecology*, **25**, 495-507.

Jakobsen, H. B., Friis, P., Nielsen, J. K., & Olsen, C. E. (1994). Emission of volatiles from flowers and leaves of *Brassica napus* in situ. *Phytochemistry*, **37**, 695-699.

Jansson, S. & Brydegaard, M. (2018). Passive kHz lidar for the quantification of insect activity and dispersal. *Animal Biotelemetry*, **6**, 1-10.

Jaquiéry, J., Broquet, T., Hirzel, A. H., Yearsley, J., & Perrin, N. (2011). Inferring landscape effects on dispersal from genetic distances: how far can we go? *Molecular Ecology*, **20**, 692-705.

Jeger, M. J. (1999). Improved understanding of dispersal in crop pest and disease management: current status and future directions. *Agricultural and Forest Meteorology*, **97**, 331-349.

Jepson, W. F. & Southwood, T. R. E. (1958) Population studies on *Oscinella frit* L. *Annals of Applied Biology*, **46**, 465-474.

Johnen A. and von Richthofen J.-S. (2013) The decision-support system proPlant expert: a computer-based tool for integrated pest management in Europe. *IOBC-WPRS Bulletin*, **96**, 99–105.

Johnen, A., Williams, I. H., Nilsson, C., Klukowski, Z., Luik, A., & Ulber, B. (2010). The proPlant decision support system: Phenological models for the major pests of oilseed rape and their key parasitoids in Europe. In: *Biocontrol-Based Integrated Management of Oilseed Rape Pests* (pp. 381-403). Springer Netherlands.

Johnson, C. G. (1957a). The distribution of insects in the air and the empirical relation of density to height. *The Journal of Animal Ecology*, **26**, 479-494.

Johnson, C. G. (1957b). The vertical distribution of aphids in the air and the temperature lapse rate. *Quarterly Journal of the Royal Meteorological Society*, **83**, 194-201.

Johnson, C. G. (1969). *Migration and dispersal of insects by flight*. Methuen & Co., London.

Johnson, C. G., Taylor, L. R., & Southwood, T. R. E. (1962). High altitude migration of Oscinella frit L.(Diptera: Chloropidae). *The Journal of Animal Ecology*, 373-383.

Johnson, D. H. (2012). Monitoring that matters. In: *Design and analysis of long-term ecological monitoring studies*. Cambridge University Press.

Jones, C. J., Isard, S. A., & Cortinas, M. R. (1999). Dispersal of synanthropic Diptera: lessons from the past and technology for the future. *Annals of the Entomological Society of America*, **92**, 829-839.

Jones, H. B., Lim, K. S., Bell, J. R., Hill, J. K., & Chapman, J. W. (2016). Quantifying interspecific variation in dispersal ability of noctuid moths using an advanced tethered flight technique. *Ecology and Evolution*, **6**, 181-190.

Jones, R. E., Gilbert, N., Guppy, M., & Nealis, V. (1980). Long-distance movement of *Pieris rapae*. *The Journal of Animal Ecology*, 629-642.

Jönsson, M., Lindkvist, A., & Anderson, P. (2005). Behavioural responses in three ichneumonid pollen beetle parasitoids to volatiles emitted from different phenological stages of oilseed rape. *Entomologia Experimentalis et Applicata*, **115**, 363-369.

Joseph, M. B. (2020). Neural hierarchical models of ecological populations. *Ecology Letters* https://doi.org/10.1111/ele.13462.

Jourdheuil, P. (1961). *Influence de quelques facteurs écologique sur les fluctuations de population d'une biocénose parasitaire: étude relative à quelques Hyménoptères (Ophioninae, Diospilinae, Euphorinae) parasites de divers Coléoptères inféodés aux Crucifères* (Doctoral dissertation, Institut national de la recherche agronomique).

Juhel, A. S., Barbu, C. M., Franck, P., Roger-Estrade, J., Butier, A., Bazot, M., & Valantin-Morison, M. (2017). Characterization of the pollen beetle, *Brassicogethes aeneus*, dispersal from woodlands to winter oilseed rape fields. *PloS One*, **12**.

Juhel, A. S., Barbu, C. M., Valantin-Morison, M., Gauffre, B., Leblois, R., Olivares, J., & Franck, P. (2019). Limited genetic structure and demographic expansion of the *Brassicogethes aeneus* populations in France and in Europe. *Pest Management Science*, **75**, 667-675.

Junjing, L., Leigang, S., & Wenjiang, H. (2019). Research progress in monitoring and forecasting of crop diseases and pests by remote sensing. *Remote Sensing Technology and Application*, **34**, 21-32.

Junk, J., Jonas, M., & Eickermann, M. (2016). Assessing meteorological key factors influencing crop invasion by pollen beetle (*Meligethes aeneus* F.)–past observations and future perspectives. *Meteorologische Zeitschrift*, 357-364.

Kaasik, R., Kovács, G., Kaart, T., Metspalu, L., Williams, I. H., & Veromann, E. (2014a). *Meligethes aeneus* oviposition preferences, larval parasitism rate and species composition of parasitoids on *Brassica nigra*, *Raphanus sativus* and *Eruca sativa* compared with on *Brassica napus*. *Biological Control*, **69**, 65-71.

Kaasik, R., Kovács, G., Toome, M., Metspalu, L., & Veromann, E. (2014b). The relative attractiveness of *Brassica napus, B. rapa, B. juncea* and *Sinapis alba* to pollen beetles. *BioControl*, **59**, 19-28.

Kaiser, C., Kristensen, M., Nauen, R., & Köhler, H. (2014). Monitoring of insecticide resistance in pollen beetle (*Meligethes aeneus* F.) populations in northern countries. In: *10th European Congress of Entomology*.

Kaiser, C., Kristensen, M., & Jensen, K. M. V. (2015). Biannual monitoring of pyrethroid and neonicotinoid susceptibility in Danish pollen beetle (*Meligethes aeneus* F.) populations. *Njf Report*.

Kaitala, V., & Ranta, E. (1998). Travelling wave dynamics and self-organization in a spatio-temporally structured population. *Ecology Letters*, **1**, 186-192.

Kaitala, V., Ranta, E. & Lindström, J. (1996). External perturbations and cyclic dynamics in stable populations. *Annales Zoologici Fennici,* **33**, 275-282

Kane, M. J., Price, N., Scotch, M., & Rabinowitz, P. (2014). Comparison of ARIMA and Random Forest time series models for prediction of avian influenza H5N1 outbreaks. *BMC Bioinformatics*, **15**, 276.

Karlson, P., & Lüscher, M. (1959). 'Pheromones': a new term for a class of biologically active substances. *Nature*, **183**, 55-56.

Kasprowicz, L., Malloch, G., Pickup, J. & Fenton, B. (2008) Spatial and temporal dynamics of *Myzus persicae* clones in fields and suction traps. *Agricultural and Forest Entomology,* **10,** 91-100.

Kassambara, A. (2019) *Package 'ggpubr': 'ggplot2' Based Publication Ready Plots*. R package version 0.2.4.

Kassambara, A. & Mundt, F. (2020). *Factoextra: Extract and Visualize the Results of Multivariate Data Analyses*. R package version 1.0.7. https://CRAN.R-project.org/package=factoextra.

Kazachkova, N., Meijer, J., & Ekbom, B. (2007). Genetic diversity in pollen beetles (*Meligethes aeneus*) in Sweden: role of spatial, temporal and insecticide resistance factors. *Agricultural and Forest Entomology*, **9**, 259-269.

Kazachkova, N., Meijer, J., & Ekbom, B. (2008). Genetic diversity in European pollen beetle, *Meligethes aeneus* (Coleoptera: Nitidulidae), populations assessed using AFLP analysis. *European Journal of Entomology*, **105**.

Kees, A. M., Hefty, A. R., Venette, R. C., Seybold, S. J., & Aukema, B. H. (2017). Flight capacity of the walnut twig beetle (Coleoptera: Scolytidae) on a laboratory flight mill. *Environmental Entomology*, **46**, 633-641.

Keitt, T. H., Bjørnstad, O. N., Dixon, P. M., & Citron-Pousty, S. (2002). Accounting for spatial pattern when modelling organism-environment interactions. *Ecography*, **25**, 616-625.

Kelly, J. F., & Stepanian, P. M. (2020). Radar Aeroecology. *Remote Sensing, **12***, 1768.

Kendall, B. E., Bjørnstad, O. N., Bascompte, J., Keitt, T. H., & Fagan, W. F. (2000). Dispersal, environmental correlation, and spatial synchrony in population dynamics. *The American Naturalist*, **155**, 628-636.

Kendon, M., McCarthy, M., Jevrejeva, S., Matthews, A., & Legg, T. (2019). State of the UK climate 2018. *International Journal of Climatology*, **39**, 1-55.

Kennedy, J. S. (1985). Migration, behavioural and ecological. In: *Migration: Mechanisms and Adaptive Significance*, (pp 5-26). *Contributions in Marine Science vol.27 (Suppl).* The University of Texas at Austin.

Kenward, H. K. (1984). Mass migration of *Meligethes aeneus* (Fabricius)(Col., Nitidulidae) at Bempton Cliffs, Yorkshire. *Entomologist's Monthly Magazine*, **120**, 236.

Khambay, B. P. S., & Jewess, P. J. (2005). Pyrethroids. *Comprehensive Molecular Insect Science*, **6**, 1-29.

Kikkawa, J. & Andersen, D. J. (1986) *Community Ecology: Pattern and Process*. Blackwell Scientific UK.

Kim, K. S., & Sappington, T. W. (2013). Population genetics strategies to characterize long-distance dispersal of insects. *Journal of Asia-Pacific Entomology*, **16**, 87-97.

Kirk, W. D. (1992). *Insects on cabbages and oilseed rape*. Richmond Publishing.

Kirkeby, C., Rydhmer, K., Cook, S. M., Strand, A., Torrance, M. T., Swain, J. L., Prangsma, J., Johnen, A., Jensen, M., Brydegaard, M. & Græsbøll, K. (2021). Advances in automatic identification of flying insects using optical sensors and machine learning. *Scientific Reports* (in press).

Kirk-Spriggs, A. H. (1996). *Pollen beetles: Coleoptera: Kateretidae and Nitidulidae: Meligethinae*. Royal Entomological Society.

Kjaer, A. (1976). *Glucosinolates in the Cruciferae* (Vol. 207). Academic Press, London.

Klick, J., Lee, J. C., Hagler, J. R., Bruck, D. J. & Yang, W. Q. (2014) Evaluating *Drosophila suzukii* immunomarking for mark–capture research. *Entomologia Experimentalis et Applicata*, **152**, 31–41.

Knape, J. (2016). Decomposing trends in Swedish bird populations using generalized additive mixed models. *Journal of Applied Ecology*, **53**, 1852-1861.

Knight, S., Kightley, S., Bingham, I., Hoad, S., Lang, B., Philpott, H., Stobart, R., Thomas, J., Barnes, A. & Ball, B. (2012). Desk study to evaluate contributory causes of the current 'yield plateau' in wheat and oilseed rape. HGCA Project Report 502.

Koenig, W. D. (1999). Spatial autocorrelation of ecological phenomena. *Trends in Ecology & Evolution*, **14**, 22-26.

Koenig, W. D. (2002). Global patterns of environmental synchrony and the Moran effect. *Ecography*, **25**, 283-288.

Kogan, M. (1998). Integrated pest management: historical perspectives and contemporary developments. *Annual Review of Entomology*, **43**, 243-270.

Kovács, G., Kaasik, R., Metspalu, L., & Veromann, E. (2013). The attractiveness of wild cruciferous plants on the key parasitoids of *Meligethes aeneus*. *IOBC-WPRS Bulletin*, **96**, 81-92.

Krenker, A., Bešter, J., & Kos, A. (2011). Introduction to the artificial neural networks. *Artificial Neural Networks: Methodological Advances and Biomedical Applications. InTech*, 1-18.

Krüger, K., Laubscher, K., Schröder, M. L., Miller, I. & Muller, L. 2014. *Aphid monitoring in the South African Potato Industry*. Potato SA Final Report.

Lack, D., & Owen, D. F. (1955). The food of the swift. *The Journal of Animal Ecology*, **24**, 120-136.

Lagos-Ortiz, K., Medina-Moreira, J., Morán-Castro, C., Campuzano, C., & Valencia-García, R. (2018). An ontology-based decision support system for insect pest control in crops. In: *International Conference on Technologies and Innovation* (pp. 3-14). Springer, Cham.

Lamb, R. J. (1989). Entomology of oilseed *Brassica* crops. *Annual Review of Entomology*, **34**, 211-229.

Lande, R., Engen, S., & Sæther, B. E. (1999). Spatial scale of population synchrony: environmental correlation versus dispersal and density regulation. *The American Naturalist*, **154**, 271-281.

Lantschner, M. V., Aukema, B. H., & Corley, J. C. (2019). Droughts drive outbreak dynamics of an invasive forest insect on an exotic host. *Forest Ecology and Management*, **433**, 762-770.

Láska, P., & Kocourek, F. (1991). Monitoring of flight activity in some crucifer-feeding pest by means of yellow water-traps. *Acta Entomologica Bohemoslovaca (CSFR)*.

Láska, P., Zelenkova, I., & Bicik, V. (1986). Colour attraction in species of the genera: *Delia* (Diptera, Anthomyidae), *Ceutorhynchus, Meligethes* and *Phyllotreta* (Coleoptera: Curculionidae, Nitidulidae, Chrysomelidae). *Acta Entomologica Bohemoslovaca*.

Leather, S. R. & Lehti, J. P. (1982). Field studies on the factors affecting the population dynamics of the bird cherry-oat aphid, *Rhopalosiphum padi* (L.) in Finland. *Ann. Agric. Fenn*., **21**, 20-31.

LeCun, Y., Bengio, Y. & Hinton, G. (2015). Deep learning. *Nature,* **521**, 436-444.

Legendre, P. (1993). Spatial autocorrelation: trouble or new paradigm?. *Ecology*, **74**, 1659-1673.

Legendre, P., Dale, M. R., Fortin, M. J., Gurevitch, J., Hohn, M., & Myers, D. (2002). The consequences of spatial structure for the design and analysis of ecological field surveys. *Ecography*, **25**, 601-615.

Legendre, P., & Fortin, M. J. (1989). Spatial pattern and ecological analysis. *Vegetatio*, **80**, 107-138.

Legg, C. J., & Nagy, L. (2006). Why most conservation monitoring is, but need not be, a waste of time. *Journal of Environmental Management*, **78**, 194-199.

Lennon, J. J. (2000). Red-shifts and red herrings in geographical ecology. *Ecography*, **23**, 101-113.

Lerch, S., Thorarinsdottir, T. L., Ravazzolo, F., & Gneiting, T. (2017). Forecaster's dilemma: Extreme events and forecast evaluation. *Statistical Science*, **32**, 106-127.

Lerin, J. (1988). Yield losses associated with 2 successive pests (*Ceuthorrhynchus napi* Gyll. and *Meligethes aeneus* F.) on winter rape (cultivar Bienvenu). *Agronomie*.

Leskinen, M., Markkula, I., Koistinen, J., Pylkkö, P., Ooperi, S., Siljamo, P., Ojanen, H., Raiskio, S. & Tiilikkala, K. (2011). Pest insect immigration warning by an atmospheric dispersion model, weather radars and traps. *Journal of Applied Entomology*, **135**, 55-67.

Lewis, T., & Taylor, L. R. (1965). Diurnal periodicity of flight by insects. *Transactions of the Royal Entomological Society of London*, **116**, 393-435.

Lewsey, J. D., & Thomson, W. M. (2004). The utility of the zero-inflated Poisson and zero-inflated negative binomial models: a case study of cross-sectional and longitudinal DMF data examining the effect of socio-economic status. *Community Dentistry and Oral Epidemiology*, **32**, 183-189.

Leybourne, D. J., Storer, K. E., Berry, P., & Ellis, S. (2021). Development of a pest threshold decision support system for minimising damage to winter wheat from wheat bulb fly, *Delia coarctata*. *bioRxiv*.

Liaw, A. & Wiener, M. (2015). *RandomForest*. R package version 4, 6–12.

Liebhold, A. M., & Gurevitch, J. (2002). Integrating the statistical analysis of spatial data in ecology. *Ecography*, **25**, 553-557.

Liebhold, A. M., Johnson, D. M., & Bjørnstad, O. N. (2006). Geographic variation in density-dependent dynamics impacts the synchronizing effect of dispersal and regional stochasticity. *Population Ecology*, **48**, 131-138.

Liebhold, A., Koenig, W. D., & Bjørnstad, O. N. (2004). Spatial synchrony in population dynamics. *Annual Reviews Ecology Evolution Systems*, **35**, 467-490.

Lim, K. S., Wolf, M., Jones, H., & Black, I. (2013). Flight mill. *Patent number: PCT/GB2014/052466*.

Lima, C. H. O., Sarmento, R. A., Rosado, J. F., Silveira, M. C. A. C., Santos, G. R., Pedro Neto, M., Erasmo, E. A. L., Nascimento, I. R. & Picanço, M. C. (2014). Efficiency and economic feasibility of pest control systems in watermelon cropping. *Journal of Economic Entomology*, **107**, 1118-1126.

Lin, X., & Zhang, D. (1999). Inference in generalized additive mixed modelsby using smoothing splines. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **61**, 381-400.

Lindblom, J., Lundström, C., Ljung, M., & Jonsson, A. (2017). Promoting sustainable intensification in precision agriculture: review of decision support systems development and strategies. *Precision Agriculture*, **18**, 309-331.

Lindelöw, Å. & Weslien, J. (1986) Sex-specific emergence of *Ips typographus* L. (Coleoptera: Scolytidae) and flight behaviour in response to pheromone sources following hibernation. *Canadian Entomologist*, **118**, 59-67.

Lindenmayer, D. B., Franklin, J. F., Lõhmus, A., Baker, S. C., Bauhus, J., Beese, W., Brodie, A., Kiehl, B., Kouki, J., Pastur, G. M. & Gustafsson, L. (2012). A major shift to the retention approach for forestry can help resolve some global forest sustainability issues. *Conservation Letters*, **5**, 421-431.

Lindenmayer, D. B., & Likens, G. E. (2010). The science and application of ecological monitoring. *Biological Conservation*, **143**, 1317-1328.

Lintner, J. A. (1882). *New principle in protection from insect attack.*

Loh, W. Y. (2011). Classification and regression trees. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, **1**, 14-23.

Longley, M., Jepson, P. C., Izquierdo, J., & Sotherton, N. (1997). Temporal and spatial changes in aphid and parasitoid populations following applications of deltamethrin in winter wheat. *Entomologia Experimentalis et Applicata*, **83**, 41-52.

López-Pantoja, G., Nevado, L. D., & Sánchez-Osorio, I. (2008). Mark-recapture estimates of the survival and recapture rates of *Cerambyx welensii* Küster (Coleoptera Cerambycidae) in a cork oak dehesa in Huelva (Spain). *Central European Journal of Biology*, **3**, 431-441.

Lövei, G. L., Stringer, I. A., Devine, C. D., & Cartellieri, M. (1997). Harmonic radar-a method using inexpensive tags to study invertebrate movement on land. *New Zealand Journal of Ecology*, 187-193.

Lovett, G. M., Burns, D. A., Driscoll, C. T., Jenkins, J. C., Mitchell, M. J., Rustad, L., Shanley, J. B., Likens, G. E. & Haeuber, R. (2007). Who needs environmental monitoring?. *Frontiers in Ecology and the Environment*, **5**, 253-260.

Lu J., Ehsani R., Shi Y., Abdulridha J. & Castro, A.I. (2017). Field detection of anthracnose crown rot in strawberry using spectroscopy technology. *Computers and Electronics in Agriculture*, **135**, 289–299. https://doi.org/10.1016/j.compag.2017.01.017.

Lüdecke D. (2018). “ggeffects: Tidy Data Frames of Marginal Effects from Regression Models.” *Journal of Open Source Software*, **3**, 772. Doi: 10.21105/joss.00772.

Lundström, C. & Lindblom, J. (2018). Considering farmers’ situated knowledge of using agricultural decision support systems (AgriDSS) to Foster farming practices: The case of CropSAT. *Agricultural Systems*, **159**, 9-20.

Lunt, D. H., Zhang, D. X., Szymura, J. M., & Hewltt, O. M. (1996). The insect cytochrome oxidase I gene: evolutionary patterns and conserved primers for phylogenetic studies. *Insect Molecular Biology*, **5**, 153-165.

Macaulay, E. D. M., Tatchell, G. M., & Taylor, L. R. (1988). The Rothamsted Insect Survey ‘12-metre' suction trap. *Bulletin of Entomological Research*, **78**, 121-128.

Magurran, A. E., Baillie, S. R., Buckland, S. T., Dick, J. M., Elston, D. A., Scott, E. M., Smith, R. I., Somerfield, P. J. & Watt, A. D. (2010). Long-term datasets in biodiversity research and monitoring: assessing change in ecological communities through time. *Trends in Ecology & Evolution*, **25**, 574-582.

Maitland, D. P., Jackson, R. L., Ladle, R. J., & Ward, P. (2002). Field considerations and problems associated with radio tracking a tropical fresh-water land crab. *Journal of Crustacean Biology*, **22**, 493-496.

Makūnas, V. (2012). *Species diversity of pollen beetles (Meligethes sl: Coleoptera, Nitidulidae) in oilseed rape and resistance of (Meligethes aeneus (F.) to pyrethroids* (Doctoral dissertation, Aleksandras Stulginskis University).

Manamperi, W. A., Chang, S. K., Ulven, C. A., & Pryor, S. W. (2010). Plastics from an improved canola protein isolate: preparation and properties. *Journal of the American Oil Chemists' Society*, **87**, 909-915.

Marczali, Z. & Nádasy, M. (2006). Wintering characteristic of the *Meligethes* species in Hungary. *Journal of Central European Agriculture*, **7**, 283-288.

Markkula, I., Leskinen, M., Pylkkö, P., Koistinen, J., Ooperi, S., Tiilikkala[1], K., Ojanen, H & Raiskio, S. (2008). Early warning system for insect migration using weather radars. *Zemdirbyste-Agriculture,* **95**, 110–115.

Martin, G., Durand, J. L., Duru, M., Gastal, F., Julier, B., Litrico, I., Louarn, G., Médiène, S., Moreau, D., Valantin-Morison, M. & Novak, S. (2020). Role of ley pastures in tomorrow's cropping systems. A review. *Agronomy for Sustainable Development*, **40**.

Martini, X., Rivera, M., Hoyte, A., Sétamou, M., & Stelinski, L. (2018). Effects of wind, temperature, and barometric pressure on Asian citrus psyllid (Hemiptera: Liviidae) flight behavior. *Journal of Economic Entomology*, **111**, 2570-2577.

Martini, X., & Stelinski, L. L. (2017). Influence of abiotic factors on flight initiation by Asian citrus psyllid (Hemiptera: Liviidae). *Environmental Entomology*, **46**, 369-375.

Mascanzoni, D., & Wallin, H. (1986). The harmonic radar: a new method of tracing insects in the field. *Ecological Entomology*, **11**, 387-390.

Mason, P. G., Olfert, O., Sluchinski, L., Weiss, R. M., Boudreault, C., Grossrieder, M., & Kuhlmann, U. (2003). Actual and potential distribution of an invasive canola pest, *Meligethes viridescens* (Coleoptera: Nitidulidae), in Canada. *The Canadian Entomologist*, **135**, 405-413.

Masters, D., & Luschi, C. (2018). Revisiting small batch training for deep neural networks. *arXiv preprint arXiv:1804.07612*.

Masurat, G. (1968). Vergleichende Untersuchungen über Beziehungen zwischen dem jährlichen Erstauftreten von Schadinsekten und Faktoren der Umwelt. *Archives of Phytopathology & Plant Protection*, **4**, 3-29.

Matsumura, K., Gaitan, C. F., Sugimoto, K., Cannon, A. J., & Hsieh, W. W. (2015). Maize yield forecasting by linear regression and artificial neural networks in Jilin, China. *The Journal of Agricultural Science*, **153**, 399-410.

Mauchline, A. L. (2003). *Behavioural and chemical ecology of* Meligethes aeneus*: effects of non-host plant volatiles* (PhD Thesis, Open University).

Mauchline, A. L., Birkett, M. A., Woodcock, C. M., Pickett, J. A., Osborne, J. L., & Powell, W. (2008). Electrophysiological and behavioural responses of the pollen beetle, *Meligethes aeneus*, to volatiles from a non-host plant, lavender, *Lavandula angustifolia* (Lamiaceae). *Arthropod-Plant Interactions*, **2**, 109-115.

Mauchline, A. L., Cook, S. M., Powell, W., & Osborne, J. L. (2013). Effects of non-host plant odour on *Meligethes aeneus* during immigration to oilseed rape. *Entomologia Experimentalis et Applicata*, **146**, 313-320.

Mauchline, A. L., Cook, S. M., Powell, W., Chapman, J. W., & Osborne, J. L. (2017). Migratory flight behaviour of the pollen beetle *Meligethes aeneus*. *Pest Management Science*.

Mauchline, A. L., Osborne, J. L., Martin, A. P., Poppy, G. M., & Powell, W. (2005). The effects of non-host plant essential oil volatiles on the behaviour of the pollen beetle *Meligethes aeneus. Entomologia Experimentalis et Applicata*, **114**, 181-188.

McBratney, A., Whelan, B., Ancev, T., & Bouma, J. (2005). Future directions of precision agriculture. *Precision Agriculture*, **6**, 7-23.

McCravy, K. W. (2018). A review of sampling and monitoring methods for beneficial arthropods in agroecosystems. *Insects*, **9**, 170.

McVean, R. I. K., Dixon, A. F. G., & Harrington, R. (1999). Causes of regional and yearly variation in pea aphid numbers in eastern England. *Journal of Applied Entomology*, **123**, 495-502.

Médiène, S., Valantin-Morison, M., Sarthou, J. P., De Tourdonnet, S., Gosme, M., Bertrand, M., Roger-Estrade, J., Aubertot, J. N., Rusch, A., Motisi, N. & Pelosi, C. (2011). Agroecosystem management and biotic interactions: a review. *Agronomy for Sustainable Development*, **31**, 491-514.

Mei, L., Guan, Z. G., Zhou, H. J., Lv, J., Zhu, Z. R., Cheng, J. A., Löfstedt, C., Svanberg, S. & Somesfalean, G. (2012). Agricultural pest monitoring using fluorescence lidar techniques. *Applied Physics B*, **106**, 733-740.

Meirmans, P. G. (2006). Using the AMOVA framework to estimate a standardized genetic differentiation measure. *Evolution*, **60**, 2399-2402.

Meloun, M., Militký, J., Hill, M., & Brereton, R. G. (2002). Crucial problems in regression modelling and their solutions. *Analyst*, **127**, 433-450.

Menze, B. H., Kelm, B. M., Masuch, R., Himmelreich, U., Bachert, P., Petrich, W., & Hamprecht, F. A. (2009). A comparison of random forest and its Gini importance with standard chemometric methods for the feature selection and classification of spectral data. *BMC Bioinformatics*, **10**, 213.

Metspalu, L., Williams, I. H., Jogar, K., Ploomi, A., Hiiesaar, K., Laeaeniste, P., Švilponis, E., Maend, M. & Luik, A. (2011). Distribution of *Meligethes aeneus* (F.) and *M. viridescens* (F.) on cruciferous plants. *Zemdirbyste*, **98**, 27-34.

Miao, J., Guo, P., Li, H., Wei, C., Liu, Q., Gong, Z., Duan, Y., Li, T., Jiang, Y., Feng, H. & Wu, Y. (2021). Low Barometric Pressure Enhances Tethered-Flight Performance and Reproductive of the Oriental Armyworm, Mythimna separata (Lepidoptera: Noctuidae). *Journal of Economic Entomology.*

Miao, L., Zheng, J. F., Cheng, Q. Q., Jia, Z. L., Wang, H. T., Liang, H. B., Zhang, H., Li, X., Zhang, J. H., Jiang, L. Y. & Qin, Q. L. (2011). Construction of a preliminary network of suction traps to monitor the migration of alate aphids in China. *Chinese Journal of Applied Entomology*, **48**, 1874-1878.

Midgarden, D. G., Youngman, R. R., & Fleischer, S. J. (1993). Spatial analysis of counts of western com rootworm (Coleoptera: Chrysomelidae) adults on yellow sticky traps in corn: geostatistics and dispersion indices. *Environmental Entomology*, **22**, 1124-1133.

Miller, H. J. (2004). Tobler's first law and spatial analysis. *Annals of the Association of American Geographers*, **94**, 284-289.

Miller, K., Huettmann, F., Norcross, B., & Lorenz, M. (2014). Multivariate random forest models of estuarine-associated fish and invertebrate communities. *Marine Ecology Progress Series*, **500**, 159-174.

Miller, M. R., Dunham, J. P., Amores, A., Cresko, W. A., & Johnson, E. A. (2007). Rapid and cost-effective polymorphism identification and genotyping using restriction site associated DNA (RAD) markers. *Genome Research*, **17**, 240-248.

Miller, W. E., & Epstein, M. E. (1986). Synchronous population fluctuations among moth species (Lepidoptera). *Environmental Entomology*, **15**, 443-447.

Milovanović, P., Kljajić, P., Popović, T., Andrić, G., & Pražić-Golić, M. (2019). Dynamic of *Brassicogethes aeneus* (F.)(Coleoptera, Nitidulidae) populations in Serbia's downriver Danube section and their susceptibility to insecticides. *Spanish Journal of Agricultural Research*, **17**, e1008.

Minter, M., Pearson, A., Lim, K. S., Wilson, K., Chapman, J. W., & Jones, C. M. (2018). The tethered flight technique as a tool for studying life-history strategies associated with migration in insects. *Ecological Entomology*, **43**, 397-411.

Mitchell, E. R., Hu, G., & Johanowicz, D. (2000). Management of diamondback moth (Lepidoptera: Plutellidae) in cabbage using collard as a trap crop. *HortScience*, **35**, 875-879.

Moeed, A. & Meads, M. J. (1987) Seasonality and density of emerging insects of a mixed lowland broadleaf-podocarp forest floor, Orongorongo Valley, New Zealand. *New Zealand Journal of Zoology*, **14**, 477-492.

Moericke, V. (1951). Eine Farbfalle zur Kontrolle des Fluges von Blattläusen, insbesondere der Pfirsichblattlaus, *Myzodes persicae* (Sulz.). *Nachr.-Bl. Dtsch. Pfl.-Schutzdienst*, **3**, 23-24.

Montgomery, G. A., Dunn, R. R., Fox, R., Jongejans, E., Leather, S. R., Saunders, M. E., Shortall C. R., Tingley M. W. & Wagner, D. L. (2019). Is the insect apocalypse upon us? How to find out. *Biological Conservation*, **241**, 108327.

Morales-Hojas, R., Gonzalez-Uriarte, A., Alvira Iraizoz, F., Jenkins, T., Alderson, L., Kruger, T., Hall, M. J., Greenslade, A., Shortall, C. R. & Bell, J. R. (2020). Population genetic structure and predominance of cyclical parthenogenesis in the bird cherry-oat aphid *Rhopalosiphum padi* in England. *Evolutionary Applications*, **13**, 1009-1025.

Moran, P. A. (1953). The statistical analysis of the Canadian lynx cycle. *Australian Journal of Zoology*, **1**, 291-298.

Morandin, L. A., Long, R. F., & Kremen, C. (2014). Hedgerows enhance beneficial insects on adjacent tomato fields in an intensive agricultural landscape. *Agriculture, Ecosystems & Environment*, **189**, 164-170.

Morton, A. C. (1982). The effects of marking and capture on recapture frequencies of butterflies. *Oecologia*, **53**, 105-110.

Moser, D., Drapela, T., Zaller, J. G., & Frank, T. (2009). Interacting effects of wind direction and resource distribution on insect pest densities. *Basic and Applied Ecology*, **10**, 208-215.

Mueller, U. G., & Wolfenbarger, L. L. (1999). AFLP genotyping and fingerprinting. *Trends in Ecology & Evolution*, **14**, 389-394.

Müller, H. J. (1941). Weitere beiträge zur biologie des rapsglanzkäfers, Meligethes aeneus F.(Ueber das winterlager und die massenbewegung im frühjahr). *Zeitschrift für Pflanzenkrankheiten (Pflanzenpathologie) und Pflanzenschutz*, **51**, 529-595.

Mumford, J. D., & Norton, G. A. (1984). Economics of decision making in pest management. *Annual review of entomology*, **29**, 157-174.

Murlis, J., Elkinton, J. S., & Carde, R. T. (1992). Odor plumes and how insects use them. *Annual Review of Entomology*, **37**, 505-532.

Musiolek, D., & Kočárek, P. (2016). Weather-dependent microhabitat use by *Tetrix tenuicornis* (Orthoptera: Tetrigidae). *The Science of Nature*, **103**, 1-10.

Muthmann, R. (2007). The use of plant protection products in the European Union. *Eurostat statistical books. European Commission, Luxembourg*.

Nathan, R. (2001). The challenges of studying dispersal. *Trends in Ecology & Evolution*, **16**, 481-483.

Nathan, R., Getz, W. M., Revilla, E., Holyoak, M., Kadmon, R., Saltz, D., & Smouse, P. E. (2008). A movement ecology paradigm for unifying organismal movement research. *Proceedings of the National Academy of Sciences*, **105**, 19052-19059.

Nauen, R. (2007). Pyrethroid resistance and its management in European populations of pollen beetles, *Meligethes aeneus*, in winter oilseed rape. *Proc XVI Int Plant Protection Cong 2007*, 522-523.

Nauen, R., Zimmer, C. T., Andrews, M., Slater, R., Bass, C., Ekbom, B., Gustafsson, G., Hansen, L. M. Kristensen, M., Zebitz, C. P. W. & Williamson, M. S. (2012). Target-site resistance to pyrethroids in European populations of pollen beetle, *Meligethes aeneus* F. *Pesticide Biochemistry and Physiology*, **103**, 173-180.

Neher, R. A., & Shraiman, B. I. (2011). Genetic draft and quasi-neutrality in large facultatively sexual populations. *Genetics*, **188**, 975-996.

Neher, R. A., Shraiman, B. I., & Fisher, D. S. (2010). Rate of adaptation in large sexual populations. *Genetics*, **184**, 467-481.

Nei, M., & Li, W. H. (1979). Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proceedings of the National Academy of Sciences*, **76**, 5269-5273.

Nestel, D., Carvalho, J., & Nemny-Lavy, E. (2004). The spatial dimension in the ecology of insect pests and its relevance to pest management. In: *Insect pest management* (pp. 45-63). Springer, Berlin, Heidelberg.

Newcastle University (2006) *The Economics of Oilseed Rape in England, 2004 Special Studies in Agricultural Economics No. 65.* Farm Business Survey School of Agriculture, Food and Rural Development University of Newcastle upon Tyne.

Nielsen, P. S. & Axelsen, J. (1988a). Developmental time and mortality of the immature stages of the pollen beetle (*Meligethes aeneus* F) under natural conditions. *Journal of Applied Entomology,* **105**, 198-204.

Nielsen, P. S. & Axelsen, J. (1988b). Spatial distribution of the pollen beetle (*Meligethes aeneus* F.) in spring sown rape (*Brassica napus* L.) in Denmark, and its importance for sampling. *Journal of Applied Entomology*, **105**, 35-40.

Nieminen, M., Leskinen, M., & Helenius, J. (2000). Doppler radar detection of exceptional mass-migration of aphids into Finland. *International Journal of Biometeorology*, **44**, 172-181.

Nilsson, C. (1985). Impact of ploughing on emergence of pollen beetle parasitoids after hibernation. *Zeitschrift für angewandte Entomologie*, **100**, 302-308.

Nilsson, C. (1987). Yield losses in summer rape caused by pollen beetles (*Meligethes* spp.). *Swedish Journal of Agricultural Research (Sweden)*.

Nilsson, C. (1988a). The pollen beetle (*Meligethes aeneus* F.) in winter and spring rape at Alnarp 1976-1978, 1: Migration and sex ratio. *Vaextskyddsnotiser (Sweden)*.

Nilsson, C. (1988b) The pollen beetle (*Meligethes aeneus*) in winter and spring rape at Alnarp 1976-1978. III. Mortality *Vaextskyddsnotiser (Sweden)*.

Nilsson, C. (1988c). The number of larval instars of *Meligethes aeneus* (F.) in southern Sweden. *Vaextskyddsnotiser (Sweden).*

Nilsson, C. & Andreasson, B. (1987). Parasitoids and predators attacking pollen beetles (*Meligethes aeneus* F.) in spring and winter rape in Southern Sweden. *IOBC-WPRS Bulletin*, **10**, 64-73.

Northing, P. & Dale, F. (2009) Extensive field based aphid monitoring as an information tool for the UK seed potato industry. *Aspects of Applied Biology*, **94**, 31-34.

OJEU (2013). Commission Implementing Regulation (EU) No 485/2013 of 24 May 2013 amending Implementing Regulation (EU) No 540/2011, as regards the conditions of approval of the active substances clothianidin, thiamethoxam and imidacloprid, and prohibiting the use and sale of seeds treated with plant protection products containing those active substances. *OJEU*, **139**, 12-14.

Økland, B., Liebhold, A. M., Bjørnstad, O. N., Erbilgin, N., & Krokene, P. (2005). Are bark beetle outbreaks less synchronous than forest Lepidoptera outbreaks? *Oecologia*, **146**, 365-372.

O'Leary, S. J., Puritz, J. B., Willis, S. C., Hollenbeck, C. M., & Portnoy, D. S. (2018). These aren't the loci you're looking for: Principles of effective SNP filtering for molecular ecologists. *Molecular Ecology*, **27**, 3193-3206.

Oliver, T. H., & Roy, D. B. (2015). The pitfalls of ecological forecasting. *Biological Journal of the Linnean Society*, **115**, 767-778.

Olfert, O., & Weiss, R. M. (2006). Impact of climate change on potential distributions and relative abundances of *Oulema melanopus, Meligethes viridescens* and *Ceutorhynchus obstrictus* in Canada. *Agriculture, Ecosystems & Environment*, **113**, 295-301.

Onstad, D. W., Bueno, A. D. F., & Favetti, B. M. (2019). Economic thresholds and sampling in integrated pest management. *The Economics of Integrated Pest Management of Insects*, 122.

Osawa, T., Mitsuhashi, H., Uematsu, Y., & Ushimaru, A. (2011). Bagging GLM: improved generalized linear model for the analysis of zero-inflated data. *Ecological Informatics*, **6**, 270-275.

Osborne, J. L., Clark, S. J., Morris, R. J., Williams, I. H., Riley, J. R., Smith, A. D., Reynolds, D. R. & Edwards, A. S. (1999). A landscape-scale study of bumble bee foraging range and constancy, using harmonic radar. *Journal of Applied Ecology*, **36**, 519-533.

Osborne, J. L., Loxdale, H. D., & Woiwod, I. P. (2002). Monitoring insect dispersal: methods and approaches. *Dispersal Ecology*, 24-49.

Osborne, J. L., & Williams, I. H. (2001). Site constancy of bumble bees in an experimentally patchy habitat. *Agriculture, Ecosystems & Environment*, **83**, 129-141.

Osborne, P. (1960). Observations on the natural enemies of *Meligethes aeneus* (F.) and *M. viridescens* (F.)[Coleoptera: Nitidulidae]. *Parasitology*, **50**, 91-110.

Osborne, P. (1965). Morphology of the immature stages of *Meligethes aeneus* (F.) and *M. viridescens* (F.)(Coleoptera, Nitidulidae). *Bulletin of Entomological Research*, **55**, 747-759.

Ouvrard, P., Hicks, D. M., Mouland, M., Nicholls, J. A., Baldock, K. C., Goddard, M. A., Kunin, W. E., Potts, S. G., Thieme, T., Veromann, E. & Stone, G. N. (2016). Molecular taxonomic analysis of the plant associations of adult pollen beetles (Nitidulidae: Meligethinae), and the population structure of *Brassicogethes aeneus*. *Genome*, **59**, 1101-1116.

Palagacheva, N., Dospatliev, L., & Dimitrov, Y. (2014). Efficient pest control of Pollen beetle (*Meligethes aeneus* F.) and possibilities for protecting the pollinators in oilseed rape agrocenosis. *International Journal of Scientific and Research Publications*, **4**, 1-4.

Palmer, J. (2020). The new science of volcanoes harnesses AI, satellites and gas sensors to forecast eruptions. *Nature*, **581**, 256-259.

Palmqvist, E., & Lundberg, P. (1998). Population extinctions in correlated environments. *Oikos*, **83**, 359-367.

Park, Y. L., Krell, R. K., Carroll, M. (2007). Theory, technology, and practice of site-specific insect pest management. *Journal of Asia-Pacific Entomology*, **10**: 89–101.

Parker, D. E., Legg, T. P., & Folland, C. K. (1992). A new daily central England temperature series, 1772–1991. *International Journal of Climatology*, **12**(4), 317-342.

Pearson, A. J., Bell, J. R., Subramanian, S., & Ouma, K. (2020). *Smart Armyworm Surveillance: Project Technical Report*.

Pedgley, D. E. (1993). Managing migratory insect pests—a review. *International Journal of Pest Management*, **39**, 3-12.

Pedigo, L. P., & Buntin, G. D. (Eds.). (1993). *Handbook of sampling methods for arthropods in agriculture*. CRC Press.

Pedigo, L. P., Hutchins, S. H., & Higley, L. G. (1986). Economic injury levels in theory and practice. *Annual Review of Entomology*, **31**, 341-368.

Pellegrino, A. C., Peñaflor, M. F. G. V., Nardi, C., Bezner-Kerr, W., Guglielmo, C. G., Bento, J. M. S., & McNeil, J. N. (2013). Weather forecasting by insects: modified sexual behaviour in response to atmospheric pressure changes. *PloS one*, **8**, e75004.

Peltonen, M., Liebhold, A. M., Bjørnstad, O. N., & Williams, D. W. (2002). Spatial synchrony in forest insect outbreaks: roles of regional stochasticity and dispersal. *Ecology*, **83**, 3120-3129.

Pérez-Rodríguez, J., Shortall, C. R. & Bell, J. R. (2015) Large-scale migration synchrony between parasitoids and their host. *Ecological Entomology,* **40**, 654-659.

Philibert, A., Desprez-Loustau, M. L., Fabre, B., Frey, P., Halkett, F., Husson, C., Lung-Escarmant, B., Marçais, B., Robin, C., Vacher, C. & Makowski, D. (2011). Predicting invasion success of forest pathogenic fungi from species traits. *Journal of Applied Ecology*, **48**, 1381-1390.

Pichler, M., Boreux, V., Klein, A. M., Schleuning, M., & Hartig, F. (2020). Machine learning algorithms to infer trait-matching and predict species interactions in ecological networks. *Methods in Ecology and Evolution*, **11**, 281-293.

Pollard, E. (1991). Synchrony of population fluctuations: the dominant influence of widespread factors on local butterfly populations. *Oikos*, **60**, 7-10.

Post, E., & Forchhammer, M. C. (2002). Synchronization of animal population dynamics by large-scale climate. *Nature*, **420**, 168-171.

Potamitis, I., Rigakis, I., & Tatlas, N. A. (2017). Automated surveillance of fruit flies. *Sensors*, **17**, 110.

Prasad, Y. & Prabhakar, M. (2012). Pest monitoring and forecasting. In: *Integrated pest management: principles and practice*. Oxfordshire, UK: Cabi, 41-57.

Preti, M., Verheggen, F., & Angeli, S. (2020). Insect pest monitoring with camera-equipped traps: strengths and limitations. *Journal of Pest Science*, 1-15.

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945-959.

Quarterman, K. D., Kilpatrick, J. W., & Mathis, W. (1954a). Fly dispersal in a rural area near Savannah, Georgia. *Journal of Economic Entomology*, **47**. 413-449.

Quarterman, K. D., Mathis, W., & Kilpatrick, J. W. (1954b). Urban fly dispersal in the area of Savannah, Georgia. *Journal of Economic Entomology*, **47**, 405-12.

R Core Team (2018). *R: A language and environment for statistical computing. R Foundation for Statistical Computing*, Vienna, Austria. URL https://www.R-project.org/.

Rackauckas, C., Ma, Y., Martensen, J., Warner, C., Zubov, K., Supekar, R., Skinner, D. & Ramadhan, A. (2020). Universal Differential Equations for Scientific Machine Learning. *arXiv preprint arXiv:2001.04385.*

Rafter, M. A., McCulloch, G. A., Daglish, G. J., Gurdasani, K., & Walter, G. H. (2018). Polyandry, genetic diversity and fecundity of emigrating beetles: understanding new foci of infestation and selection. *Journal of Pest Science*, **91**, 287-298.

Raimondo, S., Liebhold, A. M., Strazanac, J. S., & Butler, L. (2004). Population synchrony within and among Lepidoptera species in relation to weather, phylogeny, and larval phenology. *Ecological Entomology*, **29**, 96-105.

Rainey, R. C. (1962). The mechanisms of Desert Locust swarm movements and the migration of insects. In: *Proceedings of the XVth International Congress of Entomology* (Vol. 3, pp. 47-49).

Ramsden, M. W., Kendall, S. L., Ellis, S. A., & Berry, P. M. (2017). A review of economic thresholds for invertebrate pests in UK arable crops. *Crop Protection*, **96**, 30-43.

Rangapuram, S. S., Seeger, M. W., Gasthaus, J., Stella, L., Wang, Y., & Januschowski, T. (2018). Deep state space models for time series forecasting. In: *Advances in neural information processing systems* (pp. 7785-7794).

Ranta, E., Kaitala, V., Lindström, J., & Linden, H. (1995). Synchrony in population dynamics. *Proceedings of the Royal Society of London. Series B: Biological Sciences,* **262**, 113-118.

Reisig, D. D., Bacheler, J. S., Herbert, D. A., Kuhar, T., Malone, S., Philips, C., & Weisz, R. (2012). Efficacy and value of prophylactic vs. integrated pest management approaches for management of cereal leaf beetle (Coleoptera: Chrysomelidae) in wheat and ramifications for adoption by growers. *Journal of Economic Entomology*, **105**, 1612-1619.

Reppert, S. M., Guerra, P. A., & Merlin, C. (2016). Neurobiology of monarch butterfly migration. *Annual Review of Entomology,* **61**.

Reynolds, D. R., Chapman, J. W., & Harrington, R. (2006). The migration of insect vectors of plant and animal viruses. *Advances in Virus Research*, **67**, 453-517.

Reynolds, D. R., Chapman, J. W., & Stewart, A. J. (2017). Windborne migration of Auchenorrhyncha (Hemiptera) over Britain*. European Journal of Entomology,* **114***,* 554–564.

Reynolds, D. R., Riley, J. R., Armes, N. J., Cooter, R. J., Tucker, M. R., Colvin, J., Dent, D. R. & Walton, M. P. (1997). Techniques for quantifying insect migration, pp. 111–145. In: *Methods in ecological and agricultural entomology. CAB International, Wallingford, UK*.

Richards, O. W. & Waloff, N. (1954) Studies on the biology and population dynamics of British grasshoppers. *Anti-Locust Bulletin*, **17**, 1-182.

Richards, O. W. & Waloff, N. (1961) A study of the natural population of *Phytodecta olivacea* (Forster) (Coleoptera, Chrysomelidae). *Philosophical Transactions of the Royal Society B*, **244**, 205-257.

Richardson, D. M. (2008). Summary of findings from a participant country pollen beetle questionnaire. *EPPO bulletin*, **38**, 68-72.

Riggi, L. G., Gagic, V., Bommarco, R., & Ekbom, B. (2016). Insecticide resistance in pollen beetles over 7 years–a landscape approach. *Pest Management Science,* **72**, 780-786*.*

Riggi, L. G., Gagic, V., Rusch, A., Malsher, G., Ekbom, B., & Bommarco, R. (2017). Pollen beetle mortality is increased by ground-dwelling generalist predators but not landscape complexity. *Agriculture, Ecosystems & Environment*, **250**, 133-142.

Riley, J. (1995). *Adaption of harmonic radar for tracking tsetse flies*. Final Technical Report.

Riley, J. R., Reynolds, D. R., Smith, A. D., Edwards, A. S., Zhang, X. X., Cheng, X. N., Wang, H. K., Cheng, J. Y. & Zhai, B. P. (1995). Observations of the autumn migration of the rice leaf roller *Cnaphalocrocis medinalis* (Lepidoptera: Pyralidae) and other moths in eastern China. *Bulletin of Entomological Research*, **85**, 397-414.

Riley, J. R., Smith, A. D., Reynolds, D. R., Edwards, A. S., Osborne, J. L., Williams, I. H., Carreck, N. L. & Poppy, G. M. (1996). Tracking bees with harmonic radar. *Nature*, **379**, 29-30.

Ripa, J. (2000). Analysing the Moran effect and dispersal: their significance and interaction in synchronous population dynamics. *Oikos*, **89**, 175-187.

Ripley, B. D. (1996). Tree Structured Classifiers. In: *Pattern Recognition and Neural Networks* (pp. 213-242). Cambridge: Cambridge University Press.

Ripley, B. (2019). *tree: Classification and Regression Trees*. R package version 1.0-40. https://CRAN.R-project.org/package=tree.

Robertson, G. W., Griffiths, D. W., Smith, W. M., & Butcher, R. D. (1993). The application of thermal desorption-gas chromatography-mass spectrometry to the analyses of flower volatiles from five varieties of oilseed rape (*Brassica napus* spp. *oleifer*). *Phytochemical Analysis*, **4**, 152-157.

Robinson, J. (1979) The sheep headfly, *Hydrotaea irritans* (Fallén) (Diptera: Muscidae): biology of immature stages in the soil. *Bulletin of Entomological Research*, **69**, 589-598.

Rochette, N. C., & Catchen, J. M. (2017). Deriving genotypes from RAD-seq short-read data using Stacks. *Nature Protocols*, **12**, 2640-2659.

Roff, D. A. & Fairbairn, D. J. (2001). The genetic basis of dispersal and migration, and its consequences for the evolution of correlated traits. In: *Dispersal*, (pp 191-202), Oxford University Press.

Rogers, A. (1995). Population forecasting: Do simple models outperform complex models?. *Mathematical Population Studies*, **5**, 187-202.

Rola, A. C., & Pingali, P. L. (1993). *Pesticides, rice productivity, and farmers' health: an economic assessment*. IRRI CABI.

Rose, D. C., Sutherland, W. J., Parker, C., Lobley, M., Winter, M., Morris, C., Twining, S., Ffoulkes, C., Amano, T. & Dicks, L. V. (2016). Decision support tools for agriculture: Towards effective design and delivery. *Agricultural Systems*, **149**, 165-174.

Ross, K. G. (2001). How to measure dispersal: the genetic approach. The example of fire ants. In: *Dispersal, Oxford University Press, Oxford. UK*.

Rossi, V., Salinari, F., Poni, S., Caffi, T. & Bettati, T. (2014) Addressing the implementation problem in agricultural decision support systems: the example of vite.net®. *Computers and Electronics in Agriculture*, **100**, 88-99.

Roubos, C. R. & Liburd, O. E. (2010) Evaluation of Emergence Traps for Monitoring Blueberry Gall Midge (Diptera: Cecidomyiidae) Adults and Within Field Distribution of Midge Infestation. *Journal of Economic Entomology,* **103**, 1258-1267.

Roulston, T. H., & Cane, J. H. (2000). Pollen nutritional content and digestibility for animals. In: *Pollen and pollination* (pp. 187-209). Springer Vienna.

Rousset, F. (2001). Genetic approaches to the estimation of dispersal rates. In: *Dispersal* (pp 18-42) Oxford University Press.

Rowbottom, R., Carver, S., Barmuta, L. A., Weinstein, P. and Allen, G. R. (2020), How do local differences in saltmarsh ecology influence disease vector mosquito populations? *Med Vet Entomol*, **34,** 279-290. doi:10.1111/mve.12433.

Rowley, W. A., Graham, C. L., & Williams, R. E. (1968). A flight mill system for the laboratory study of mosquito flight. *Annals of the Entomological Society of America*, **61**, 1507-1514.

Rowley, C., Pope, T. W., Cherrill, A., Leather, S. R., Fernández-Grandon, G. M., & Hall, D. R. (2017). Development and optimisation of a sex pheromone lure for monitoring populations of saddle gall midge, *Haplodiplosis marginata*. *Entomologia Experimentalis et Applicata*, **163**, 82-92.

Royama, T. (2005). Moran effect on nonlinear population processes. *Ecological Monographs*, **75**, 277-293.

Royama, T. (2012). *Analytical population dynamics* (Vol. 10). Springer Science & Business Media.

Rstudio Team (2015). *Rstudio: Integrated Development for R*. Rstudio, Inc., Boston, MA URL http://www.rstudio.com/.

Rubink, W. L. (1988). A photographic technique for producing high quality insect tags for mark/release/recapture studies. Técnica fotográfica para producir etiquetas de alta calidad para estudios de marcaje/liberación y recaptura. *Entomological News*, **99**, 167-171.

Rusch, A., Suchail, S., Valantin-Morison, M., Sarthou, J. P., & Roger-Estrade, J. (2013). Nutritional state of the pollen beetle parasitoid *Tersilochus heterocerus* foraging in the field. *BioControl*, **58**, 17-26.

Rusch, A., Valantin-Morison, M., Roger-Estrade, J., & Sarthou, J. P. (2012). Local and landscape determinants of pollen beetle abundance in overwintering habitats. *Agricultural and Forest Entomology*, **14**, 37-47.

Rusch, A., Valantin-Morison, M., Sarthou, J. P., & Roger-Estrade, J. (2011). Multi-scale effects of landscape complexity and crop management on pollen beetle parasitism rate. *Landscape Ecology*, **26**, 473-486.

Rusch, A., Valantin-Morison, M., Sarthou, J. P., & Roger-Estrade, J. (2013). Effect of crop management and landscape context on insect pest populations and crop damage. *Agriculture, Ecosystems & Environment*, **166**, 118-125.

Ruther, J., & Thiemann, K. (1997). Response of the pollen beetle *Meligethes aeneus* to volatiles emitted by intact plants and conspecifics. *Entomologia Experimentalis et Applicata*, **84**, 183-188.

Ryan, M. (2019). Agricultural Big Data Analytics and the Ethics of Power. *Journal of Agricultural and Environmental Ethics*, 1-21.

Sakamoto, Y., Ishiguro, M., & Kitagawa, G. (1986). *Akaike information criterion statistics*. Dordrecht, The Netherlands.

Sanders, C. J., Shortall, C. R., Gubbins, S., Burgin, L., Gloster, J., Harrington, R., Reynolds, D. R., Mellor, P. S. & Carpenter, S. (2011). Influence of season and meteorological parameters on flight activity of *Culicoides* biting midges. *Journal of Applied Ecology*, **48**, 1355-1364.

Sandrini Moraes, F., Edson Nava, D., Scheunemann, T., & Santos da Rosa, V. (2019). Development of an optoelectronic sensor for detecting and classifying fruit fly (diptera: Tephritidae) for use in real-time intelligent traps. *Sensors*, **19**, 1254.

Santoiemma, G., Fioretto, D., Corcos, D., Mori, N., & Marini, L. (2019). Spatial synchrony in *Drosophila suzukii* population dynamics along elevational gradients. *Ecological Entomology*, **44**, 182-189.

Saunders, M. E. (2019). No Simple Answers for Insect Conservation: Media hype has missed the biggest concern that ecologists and entomologists have about six-legged life: how little we know about it. *American Scientist*, **107**, 148.

Schaffer, W. M. (1985) Order and chaos in ecological systems. *Ecology*, **66**, 93-106.

Schaffers, A. P., Raemakers, I. P. & Sýkora, K. V. (2012) Successful overwintering of arthropods in roadside verges. *Journal of Insect Conservation*, **16,** 511-522.

Scherney, F. (1953). Zur biologie der an Raps vorkommenden Meligethesarten. *Z. Pflanzenbau Pflanzenschutz*, **48**, 154-176.

Schmidt, N. P., O'Neal, M. E., Anderson, P. F., Lagos, D., Voegtlin, D., Bailey, W., Caragea, P., Cullen, E., DiFonzo, C., Elliott, K. & Gratton, C. (2012). Spatial distribution of *Aphis glycines* (Hemiptera: Aphididae): a summary of the suction trap network. *Journal of Economic Entomology*, **105**, 259-271.

Schneider, F. (1962). Dispersal and migration. *Annual Review of Entomology*, **7**, 223-242.

Sciarretta, A., & Trematerra, P. (2014). Geostatistical tools for the study of insect spatial distribution: practical implications in the integrated management of orchard and vineyard pests. *Plant Protection Science*, **50**, 97-110.

Scott, C. & Bilsborrow, P. E. (2019). The impact of the EU neonicotinoid seed-dressing ban on oilseed rape production in England. *Pest Management Science*, **75**, 125-133.

Seddon, A. M. (1986) Abundance and life-history of four species of terrestrial Chironomidae (Dipt.) from deciduous woodland soil in south east England. *Entomologist's Monthly Magazine*, **122**, 219-228.

Šedivý, J., & Kocourek, F. (1994). Flight activity of winter rape pests. *Journal of Applied Entomology*, **117**, 400-407.

Seimandi Corda, G. (2018). Finding new targets to screen oilseed rape (*Brassica napus*) resistance to pollen beetle (*Brassicogethes aeneus*): from metabolomics to the field (Doctoral dissertation).

Seimandi Corda, G., Leblanc, M., Faure, S., & Cortesero, A. M. (2018). Impact of flower rewards on phytophagous insects: importance of pollen and nectar for the development of the pollen beetle (*Brassicogethes aeneus*). *Arthropod-Plant Interactions*, **12**, 779-785.

Seimandi Corda, G., Renaud, D., Escande, L., Larièpe, A., Ollivier, J., Faure, S., & Cortesero, A. M. (2019). Screening the variability in oilseed rape resistance to pollen beetle attacks in the field and assessment of biochemical biomarkers. *Journal of Pest Science*, **92**, 895-908.

Shabani, F., Kumar, L., & Ahmadi, M. (2016). A comparison of absolute performance of different correlative and mechanistic species distribution models in an independent area. *Ecology and Evolution*, **6**, 5973-5986.

Shaw, R. (1981) Strange attractors, chaotic behaviour, and information flow. *Zeitschrift für Naturforschung* **36a**, 80-112.

Sheppard, L. W., Bell, J. R., Harrington, R., & Reuman, D. C. (2016). Changes in large-scale climate alter spatial synchrony of aphid pests. *Nature Climate Change*, **6**, 610-613.

Shoemaker, K. T., & Akçakaya, H. R. (2015). Inferring the nature of anthropogenic threats from long-term abundance records. *Conservation Biology*, **29**, 238-249.

Shortall, C. R. (2013) *Not Just Aphids: Uses of the Rothamsted Insect Survey Sample Collection* [Poster] Conference: NJF Seminar 468, Kristianstad.

Shortall, C. R., Moore, A., Smith, E., Hall, M. J., Woiwod, I. P., & Harrington, R. (2009). Long-term changes in the abundance of flying insects. *Insect Conservation and Diversity*, **2**, 251-260.

Showers, W. B. (1997). Migratory ecology of the black cutworm. *Annual Review of Entomology*, **42**, 393-425.

Simpson, G. L. (2018). Modelling palaeoecological time series using generalised additive models. *Frontiers in Ecology and Evolution*, **6**, 149.

Skellern, M. P., & Cook, S. M. (2018). Prospects for improved off-crop habitat management for pollen beetle control in oilseed rape. *Arthropod-Plant Interactions*, **12**, 849-866.

Skellern, M. P., Welham, S. J., Watts, N. P., & Cook, S. M. (2017). Meteorological and landscape influences on pollen beetle immigration into oilseed rape crops. *Agriculture, Ecosystems & Environment*, **241**, 150-159.

Slater, R., Ellis, S., Genay, J. P., Heimbach, U., Huart, G., Sarazin, M., Longhurst, C., Müller, A., Nauen, R., Rison, J. L. & Robin, F. (2011). Pyrethroid resistance monitoring in European populations of pollen beetle (*Meligethes* spp.): a coordinated approach through the Insecticide Resistance Action Committee (IRAC). *Pest Management Science*, **67**, 633-638.

Slatkin, M. (1987). Gene flow and the geographic structure of natural populations. *Science*, **236**, 787-792.

Smart, L. E., Blight, M. M., & Ryan, J. (1995). Response of pollen beetles, *Meligethes* spp., to volatiles from *Brassica napus*. In: *Proceedings, GCIRC 9th International Rapeseed Congress* (Vol. 3, pp. 1040-1042).

Smart, L. E., & Blight, M. M. (2000). Response of the pollen beetle, *Meligethes aeneus*, to traps baited with volatiles from oilseed rape, *Brassica napus*. *Journal of Chemical Ecology*, **26**, 1051-1064.

Smith, S. K., Tayman, J., & Swanson, D. A. (2006*). State and local population projections: Methodology and analysis*. Springer, Boston, MA

Solé, R. V. & Bascompte, J. (1995) Measuring chaos from spatial information. *Journal of Theoretical Biology,* **175**, 139-147.

Song, Y. Y., & Ying, L. U. (2015). Decision tree methods: applications for classification and prediction. *Shanghai Archives of Psychiatry*, **27**, 130.

Southwood, T. R. E. (1962). Migration of terrestrial arthropods in relation to habitat. *Biological Reviews*, **37**, 171-211.

Southwood, T. R. E. & Henderson, P. A. (2000) Observational and experimental methods for the estimation of natality, mortality, and dispersal. In: *Ecological Methods 3rd Edition*, Blackwell, Oxford.

Southwood, T. R. E., & Henderson, P. A. (2009). *Ecological methods*. John Wiley & Sons.

Southwood, T. R. E. & Jepson, W. F. (1962a) The productivity of grasslands in England for *Oscinella frit* (L.) (Chloropidae) and other stem-boring diptera. *Bulletin of Entomological Research*, **53**, 395-407.

Southwood, T. R. E. & Jepson, W. F. (1962b) Studies on the populations of *Oscinella frit* L. (Dipt. Chloropidae) in the oat crop. *Journal of Animal Ecology*, **31**, 481-495.

Southwood, T. R. E. & Siddorn, J. W. (1965) Temperature beneath insect emergence traps of various types. *Journal of Animal Ecology*, **34**, 581-585.

Spiller, K. J., & Dettmers, R. (2019). Evidence for multiple drivers of aerial insectivore declines in North America. *The Condor: Ornithological Applications*.

Spitzer, T., Bílovský, J., & Matušinsky, P. (2020). Changes in resistance development in pollen beetle (*Brassicogethes aeneus* F.) to lambda-cyhalothrin, etofenprox, chlorpyrifos-ethyl, and thiacloprid in the Czech Republic during 2013–2017. *Crop Protection*, **135**, 105224.

Stechmann, D. H., & Schütte, F. (1976). Zur Ausbreitung des Rapsglanzkäfers (Meligethes aeneus F.; Col., Nitidulidae) vor der Überwinterung. *Anzeiger für Schädlingskunde, Pflanzenschutz, Umweltschutz*, **49**, 183-188.

Stepanian, P. M., Horton, K. G., Melnikov, V. M., Zrnić, D. S., & Gauthreaux Jr, S. A. (2016). Dual-polarization radar products for biological applications. *Ecosphere*, **7**, e01539.

Stephenson, D. B., Casati, B., Ferro, C. A. T., & Wilson, C. A. (2008). The extreme dependency score: a non-vanishing measure for forecasts of rare events. *Meteorological Applications*, **15**, 41-50.

Storkey, J., Macdonald, A. J., Bell, J. R., Clark, I. M., Gregory, A. S., Hawkins, N. J., Hirsch, P. R., Todman, L. C. & Whitmore, A. P. (2016). The unique contribution of Rothamsted to ecological research at large temporal scales. *Advances in Ecological Research*, **55**, 3-42. Academic Press.

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

Strand, M. R., Roitberg, B. D., & Papaj, D. R. (1990). Acridine orange: a potentially useful internal marker of Hymenoptera and Diptera. *Journal of the Kansas Entomological Society*, **63**, 634-637.

Stratonovitch, P., Elias, J., Denholm, I., Slater, R., & Semenov, M. A. (2014). An individual-based model of the evolution of pesticide resistance in heterogeneous environments: control of *Meligethes aeneus* population in oilseed rape crops. *PloS One*, **9**, e115631.

Strobl, C., Boulesteix, A. L., Zeileis, A., & Hothorn, T. (2007). Bias in random forest variable importance measures: Illustrations, sources and a solution. *BMC Bioinformatics*, **8**, 25.

Strong Jr. D. R. (1979) Biogeographic dynamics of insect-host plant communities. *Annual Review of Entomology,* **24**, 89-119.

Sutcliffe, O. L., Thomas, C. D., & Moss, D. (1996). Spatial synchrony and asynchrony in butterfly population dynamics. *Journal of Animal Ecology*, **65**, 85-95.

Sutter, L., Amato, M., Jeanneret, P., & Albrecht, M. (2018). Overwintering of pollen beetles and their predators in oilseed rape and semi-natural habitats. *Agriculture, Ecosystems & Environment*, **265**, 275-281.

Szentkirályi, F. (2002). Fifty-year-long insect survey in Hungary: T. Jermy's contributions to light-trapping. *Acta Zoologica Academiae Scientiarum Hungaricae*, **48**, 85-105.

Taimr, L., Sedivy, J., Bergmannova, E., & Hanker, I. (1967). *Further experience obtained in studies on dispersal flights of* Meligethes aeneus *f., marked with $ sup 32$ p (Coleoptera)*. Central Research Inst. for Plant Production, Prague.

Tan, S., Sim, K. C., & Gales, M. (2015). Improving the interpretability of deep neural networks with stimulated learning. In: *2015 IEEE Workshop on Automatic Speech Recognition and Understanding (ASRU)* (pp. 617-623). IEEE.

Tang, Y., Horikoshi, M., & Li, W. (2016). Ggfortify: Unified interface to visualize statistical results of popular R packages. *R J.*, **8**, 474.

Tansey, J. A., Dosdall, L. M., Keddie, A., & Olfert, O. (2010). Flight activity and dispersal of the cabbage seedpod weevil (Coleoptera: Curculionidae) are related to atmospheric conditions. *Environmental Entomology*, **39**, 1092-1100.

Tatchell, G. M. (1983). Compensation in spring-sown oil-seed rape (*Brassica napus* L.) plants in response to injury to their flower buds and pods. *The Journal of Agricultural Science*, **101**, 565-573.

Taylor, L. R. (1960). The distribution of insects at low levels in the air. *The Journal of Animal Ecology*, **29**, 45-63.

Taylor, L. R. (1962). The absolute efficiency of insect suction traps. *Annals of Applied Biology*, **50**, 405-421.

Taylor, L. R. (1973). Monitoring change in the distribution and abundance of insects. *Rothamsted Experimental Station Report for 1973*, **2**, 202-239.

Taylor, L. R. (1986). Synoptic dynamics, migration and the Rothamsted insect survey: Presidential address to the British Ecological Society, December 1984. *Journal of Animal Ecology*, **55**, 1-38.

Taylor, R. J. (2013). *Predation*. Springer Science & Business Media.

Templeton, A. R., & Read, B. (1994). Inbreeding: one word, several meanings, much confusion. pp. 91-105 In: *Conservation Genetics*. Birkhäuser, Basel.

Teng, P. S. (1994). Integrated pest management in rice. *Experimental Agriculture*, **30**, 115-137.

Thackray, D. J., Diggle, A. J., & Jones, R. A. C. (2009). BYDV PREDICTOR: a simulation model to predict aphid arrival, epidemics of Barley yellow dwarf virus and yield losses in wheat crops in a Mediterranean-type environment. *Plant Pathology*, **58**, 186-202.

Therneau, T. & Atkinson, B. (2018). *Rpart: recursive partitioning and regression trees*. R package version 4, 1–13. https://CRAN.R-project.org/package=rpart.

Thieme, T., Gloyna, K., Drbal, U., & Zellner, M. (2010a). How to determine the susceptibility of *Meligethes aeneus* to neonicotinoids?. *Journal of Applied Entomology*, **134**, 55-60.

Thieme, T., Heimbach, U. & Müller, A. (2010b). Chemical control of insect pests and insecticide resistance in oilseed rape. In: *Biocontrol-Based Integrated Management of Oilseed Rape Pests* (pp. 215-244). Springer Netherlands.

Thies, C., Steffan-Dewenter, I. & Tscharntke, T. (2008). Interannual landscape changes influence plant-herbivore-parasitoid interactions. *Agriculture, Ecosystems and Environment*, **125**, 266-268.

Thomas, C. D. (1991). Spatial and temporal variability in a butterfly population. *Oecologia*, **87**, 577-580.

Thöming, G., Solhaug, K. A., & Norli, H. R. (2020). Kairomone-assisted trap cropping for protecting spring oilseed rape (*Brassica napus*) from pollen beetles (Coleoptera: Nitidulidae). *Pest Management Science*. **76**, 3253-3263.

Thompson, K. F. & W. G. Hughes, 1986. Breeding and varieties. In: *Oilseed Rape* (pp. 32–82). Collins, London.

Tilman, D., & Kareiva, P. (Eds.). (2018). *Spatial ecology: the role of space in population dynamics and interspecific interactions (MPB-30)* (Vol. 89). Princeton University Press.

Tobler, W. R. (1970). A computer movie simulating urban growth in the Detroit region. *Economic Geography*, **46**, 234-240.

Tölle, M. L. (2014). *Factors regulating the population dynamics and damage potential of pollen beetle (*Meligethes aeneus *F.) on crops of oilseed rape* (Doctoral dissertation, Niedersächsische Staats-und Universitätsbibliothek Göttingen).

Tollsten, L., & Bergström, G. (1988). Headspace volatiles of whole plants and macerated plant parts of Brassica and Sinapis. *Phytochemistry*, **27**, 4013-4018.

Tu, W., & Liu, H. (2014). Zero-inflated data. In: *Wiley StatsRef: statistics reference online*, 1-7.

Tu, X., Li, Z., Wang, J., Huang, X., Yang, J., Fan, C. & Zhang, Z. (2014). Improving the degree-day model for forecasting *Locusta migratoria manilensis* (Meyen)(Orthoptera: Acridoidea). *PloS One*, **9**.

Tulisalo, U., & Wuori, T. (1986). Blossom beetle (*Meligethes aeneus* Fab.) as a yield factor in turnip rape (*Brassica campestris* L.). *Journal of Agricultural Science in Finland (Finland)*.

Tuovinen, T., & Parikka, P. (1996). Monitoring strawberry pests and diseases: practical applications for decision making. In: *III International Strawberry Symposium 439* (pp. 931-940).

Turchin, P. & Taylor, A. D. (1992) Complex dynamics in ecological time series. *Ecology* **73**, 289-305.

Turner, R. J. & Bowden, J. (1983) An evaluation of the use of elytra and bodies in X-ray energy-dispersive spectroscopic studies of the red turnip beetle, Entomoscelis Americana (Coleoptera:Chrysomelidae). *The Canadian Entomologist*, **112**, 609-614.

Tylianakis, J. M., Didham, R. K., Bascompte, J. and Wardle, D. A., 2008. Global change and species interactions in terrestrial ecosystems. *Ecology Letters*, **11**, 1351-1363.

Ulber, B., Fischer, K., Klukowski, Z., Luik, A., Veromann, E., Nilsson, C., Ahman, B., Williams, I. H., Ferguson, A. W., Piper, R. & Barari, H. (2006). Identity of parasitoids and their potential for biocontrol of oilseed rape pests in Europe. In: *Proc. Symp. on Integrated Pest Management in Oilseed Rape* (pp. 3-5).

Ulber, B., Williams, I. H., Klukowski, Z., Luik, A., & Nilsson, C. (2010). Parasitoids of oilseed rape pests in Europe: key species for conservation biocontrol. In: *Biocontrol-based integrated management of oilseed rape pests* (pp. 45-76). Springer, Dordrecht.

Urquhart, F. A. & Urquhart, N. R. (1979). Vernal migration of the monarch butterfly, Danaus plexippus (Lepidoptera:Danaidae) in North America from the overwintering site in the neo-volcanic plateau of Mexico. *The Canadian Entomologist*, **111**, 41-47.

Valantin-Morison, M., Meynard, J. M., & Doré, T. (2007). Effects of crop management and surrounding field environment on insect incidence in organic winter oilseed rape (*Brassica napus* L.). *Crop Protection*, **26**, 1108-1120.

Valantin-Morison, M., & Pinochet, X. (2010). Concevoir des itinéraires techniques intégrés pour le colza d'hiver: les acquis d’aujourd’hui et les défis de demain. *Innovations Agronomiques*, **8**, 35-55.

Van Doren, B. M., & Horton, K. G. (2018). A continental system for forecasting bird migration. *Science,* **361**, 1115-1118.

Van Dyck, H., Van Strien, A. J., Maes, D., & Van Swaay, C. A. (2009). Declines in common, widespread butterflies in a landscape under intense human use. *Conservation Biology*, **23**, 957-965.

van Rij J., Wieling M., Baayen R. & van Rijn H. (2017). *itsadug: Interpreting Time Series and Autocorrelated Data Using GAMMs*. R package version 2.3.

Venables W. N. & Ripley B. D. (2002). *Modern Applied Statistics with S*, Fourth edition. Springer, New York. ISBN 0-387-95457-0, http://www.stats.ox.ac.uk/pub/MASS4/.

Veran, S., Simpson, S. J., Sword, G. A., Deveson, E., Piry, S., Hines, J. E., & Berthier, K. (2015). Modeling spatiotemporal dynamics of outbreaking species: influence of environment and migration in a locust. *Ecology*, **96**, 737-748.

Veromann, E., Metspalu, L., Williams, I. H., Hiiesaar, K., Mand, M., Kaasik, R., Kovacs, G., Jogar, K., Svilponis, E., Kivimagi, I. & Ploomi, A. (2012). Relative attractiveness of *Brassica napus*, *Brassica nigra, Eruca sativa* and *Raphanus sativus* for pollen beetle (*Meligethes aeneus*) and their potential for use in trap cropping. *Arthropod-Plant Interactions*, **6**, 385-394.

Vinatier, F., Gosme, M., & Valantin-Morison, M. (2012). A tool for testing integrated pest management strategies on a tritrophic system involving pollen beetle, its parasitoid and oilseed rape at the landscape scale. *Landscape Ecology*, **27**, 1421-1433.

Vontas J., Ranson H., Williamson M. S. (2010) Addendum: Pyrethroid Insecticides and Resistance Mechanisms In: *Insect Control, Biological and Synthetic Agents*. (pp. 30–34). Elsevier; Oxford.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., Lee, T. V. D., Hornes, M., Friters, A., Pot, J., Paleman, J., Kuiper, M. & Zabeau, M. (1995). AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research*, **23**, 4407-4414.

Walter, J. A., Sheppard, L. W., Anderson, T. L., Kastens, J. H., Bjørnstad, O. N., Liebhold, A. M., & Reuman, D. C. (2017). The geography of spatial synchrony. *Ecology Letters*, **20**, 801-814.

Walters, K. F., Young, J. E., Kromp, B., & Cox, P. D. (2003). Management of oilseed rape pests. In: *Biological Control of Pests of Oilseed Rape*; Alford, DV, Ed.; Blackwell Science: Oxford, UK, 43-71.

Wang, E., Cook, D., & Hyndman, R. J. (2019). A new tidy data structure to support exploration and modeling of temporal data. *Journal of Computational and Graphical Statistics*, 1-13.

Ward, E. J., Holmes, E. E., Thorson, J. T., & Collen, B. (2014). Complexity is costly: a meta-analysis of parametric and non-parametric methods for short-term population forecasting. *Oikos*, **123**, 652-661.

Waser, P. M. & Strobeck, C. (1998). Genetic signatures of interpopulation dispersal. *Trends in Ecology and Evolution,* **13**, 43-44.

Wasmann, E. (1926). *Versuche über den Farbensinn des Rapskäfers (*Meligethes aeneus *L.)*.

Way, M. J., Murdie, G. & Galley, D. J. (1969) Experiments on integration of chemical and biological control of aphids on brussels sprouts. *Annals of Applied Biology*, **63**, 459-475.

Webb, R. E. & Smith, F. F. (1980). Greenhouse whitefly control of an integrated regimen based on adult trapping and nymphal parasitism. *Bulletin W. Palaearctic Reg. Sec. Int. Union Biol. Sci*. **3**: 235–236.

Węgorek, P., Mrówczyński, M., & Zamojska, J. (2009). Resistance of pollen beetle (*Meligethes aeneus* F.) to selected active substances of insecticides in Poland. *Journal of Plant Protection Research*, **49**, 119-128.

Wegorek, W. (1965) In: FAO 1967. Cited in IRAC 2006.

Weidel, H. (2008). *Die Verteilung des Aeroplanktons über Schleswig-Holstein* (Doctoral dissertation, Christian-Albrechts Universität Kiel).

Weightman, R., Gladders, P., & Berry, P. (2010). Oilseed rape. In: *Energy crops* (pp. 116-147). The Royal Society of Chemistry UK.

Weir, B. S. & Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution*, **38**, 1358-1370.

Welsh, A. H., Cunningham, R. B., Donnelly, C. F., & Lindenmayer, D. B. (1996). Modelling the abundance of rare species: statistical models for counts with extra zeros. *Ecological Modelling*, **88**, 297-308.

Westbrook, J. K., Eyster, R. S., Wolf, W. W., Lingren, P. D., & Raulston, J. R. (1995). Migration pathways of corn earworm (Lepidoptera: Noctuidae) indicated by tetroon trajectories. *Agricultural and Forest Meteorology*, **73**, 67-87.

Wheeler, D. L., Scott, J., Dung, J. K. S., & Johnson, D. A. (2019). Evidence of a trans-kingdom plant disease complex between a fungus and plant-parasitic nematodes. *PloS One*, **14**.

White, S. M., Sanders, C. J., Shortall, C. R., & Purse, B. V. (2017). Mechanistic model for predicting the seasonal abundance of *Culicoides* biting midges and the impacts of insecticide control. *Parasites & Vectors*, **10**, 162.

Whitehorn, P. R., O'Connor, S., Wackers, F. L., & Goulson, D. (2012). Neonicotinoid pesticide reduces bumble bee colony growth and queen production. *Science*, **336**, 351-352.

Wickham H. (2016). *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. ISBN 978-3-319-24277-4, https://ggplot2.tidyverse.org.

Wickham, H. & Henry, L. (2020). *Tidyr: Tidy Messy Data*. R package version 1.0.2. https://CRAN.R-project.org/package=tidyr.

Williams, C. B. (1957). Insect migration. *Annual Review of Entomology*, **2**, 163-180.

Williams, D. T., Straw, N., Townsend, M., Wilkinson, A. S., & Mullins, A. (2013). Monitoring oak processionary moth *Thaumetopoea processionea* L. using pheromone traps: the influence of pheromone lure source, trap design and height above the ground on capture rates. *Agricultural and Forest Entomology*, **15**, 126-134.

Williams, D. W., & Liebhold, A. M. (2000). Spatial synchrony of spruce budworm outbreaks in eastern North America. *Ecology*, **81**, 2753-2766.

Williams, I. H. (2006). Integrating parasitoids into management of pollen beetle on oilseed rape. *Agronomy Research*, **4**(special issue), 465-470.

Williams, I. H. (2010). The major insect pests of oilseed rape in Europe and their management: an overview. In: *Biocontrol-based integrated management of oilseed rape pests* (pp. 1-43). Springer Netherlands.

Williams, I. H., Büchi, R., & Ulber, B. (2003). Sampling, trapping and rearing oilseed rape pests and their parasitoids. In: *Biocontrol of Oilseed Rape Pests*, 145-160.

Williams, I. H., & Cook, S. M. (2010). Crop location by oilseed rape pests and host location by their parasitoids. In: *Biocontrol-Based Integrated Management of Oilseed Rape Pests* (pp. 215-244). Springer Netherlands.

Williams, I. H., & Ferguson, A. W. (2010). Spatio-temporal distributions of pests and their parasitoids on the oilseed rape crop. In: *Biocontrol-Based Integrated Management of Oilseed Rape Pests* (pp. 245-272). Springer Netherlands.

Williams, I. H., Ferguson, A. W., Kruus, M., Veromann, E., & Warner, D. J. (2010). Ground beetles as predators of oilseed rape pests: incidence, spatio-temporal distributions and feeding. In: *Biocontrol-based integrated management of oilseed rape pests* (pp. 115-149). Springer, Dordrecht.

Williams, I. H., Frearson, D., Barari, H., & McCartney, A. (2007). Migration to and dispersal from oilseed rape by the pollen beetle, *Meligethes aeneus*, in relation to wind direction. *Agricultural and Forest Entomology*, **9**, 279-286.

Williams, I. H., & Free, J. B. (1978). The feeding and mating behaviour of pollen beetles (*Meligethes aeneus* Fab.) and seed weevils (*Ceutorhynchus assimilis* Payk.) on oil-seed rape (*Brassica napus* L.). *The Journal of Agricultural Science*, **91**, 453-459.

Williams, I. H., & Free, J. B. (1979). Compensation of oil-seed rape (*Brassica napus* L.) plants after damage to their buds and pods. *The Journal of Agricultural Science*, **92**, 53-59.

Williams, J. N., Seo, C., Thorne, J., Nelson, J. K., Erwin, S., O'Brien, J. M., & Schwartz, M. W. (2009). Using species distribution models to predict new occurrences for rare plants. *Diversity and Distributions*, **15**, 565-576.

Williamson, B., Lawson, H. M., Woodford, J. A. T., Hargreaves, A. J., Wiseman, J. S. & Gordon, S. C. (1979) Vigour control, an integrated approach to cane, pest and disease management in red raspberry (*Rubus idaeus*). *Annals of Applied Biology*, **92**, 359-368.

Willis, K. J., Araújo, M. B., Bennett, K. D., Figueroa-Rangel, B., Froyd, C. A., & Myers, N. (2007). How can a knowledge of the past help to conserve the future? Biodiversity conservation and the relevance of long-term ecological studies. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **362**, 175-187.

Wilson, L., & Barnett, W. (1983). Degree-days: an aid in crop and pest management. *California Agriculture*, **37**, 4-7.

Winder, L. (2004). Marking by abrasion or branding and recapturing carabid beetles in studies of their movement. International Journal of Pest Management, **50**, 161-164.

Winfield, A. L. (1992). Management of oilseed rape pests in Europe. *Agricultural Zoology Reviews*, **5**, 51-95.

Wittkop, B., Snowdon, R. J., & Friedt, W. (2009). Status and perspectives of breeding for enhanced yield and quality of oilseed crops for Europe. *Euphytica*, **170**, 131-140.

Wolfenbarger, D. O. (1946). Dispersion of small organisms. Distance dispersion rates of bacteria, spores, seeds, pollen, and insects; incidence rates of diseases and injuries. *The American Midland Naturalist*, **35**, 1-152.

Wolff, M. & Krausse, A. (1926) Contribution to the knowledge of the biology of the pests of the oil-bearing crops, especially of the share in the damage by *Ceuthorrhynchus assimilis* falsely ascribed to *Meligethes aeneus*. *Arch. Naturgesch*., **91**, 1-4.

Wood, C. R., Reynolds, D. R., Wells, P. M., Barlow, J. F., Woiwod, I. P., & Chapman, J. W. (2009). Flight periodicity and the vertical distribution of high-altitude moth migration over southern Britain. *Bulletin of Entomological Research*, **99**, 525-535.

Wood, S. N. (2003). Thin plate regression splines. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **65**, 95-114.

Wood, S. N. (2011). Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology),* **73**, 3-36.

Wood, S. N. (2017). *Generalized additive models: an introduction with R*. Chapman and Hall/CRC.

Woodford, J. A. T. (1973) The climate within a large aphid-proof field cage. *Entomologia Experimentalis et Applicata*, **16**, 313-321.

Wool, D. (2002). Herbivore abundance is independent of weather? A 20-year study of a galling aphid Baizongia pistaciae (Homoptera: Aphidoidea). *Population Ecology*, **44**, 0281-0291.

Wratten, S. D. (2012). *Video techniques in animal ecology and behaviour*. Springer Science & Business Media.

Wright, S. (1943). Isolation by distance. *Genetics*, **28**, 114.

Wright, S. (1950). Genetical structure of populations. *Nature*, **166**, 247-249.

Wu, J., Aluko, R. E., & Muir, A. D. (2009). Production of angiotensin I-converting enzyme inhibitory peptides from defatted canola meal. *Bioresource Technology*, **100**, 5283-5287.

Wyatt, T. D. (2009). Fifty years of pheromones. *Nature*, **457**, 262-263.

Yadouleton, A. W. M., Asidi, A., Djouaka, R. F., Braïma, J., Agossou, C. D., & Akogbeto, M. C. (2009). Development of vegetable farming: a cause of the emergence of insecticide resistance in populations of *Anopheles gambiae* in urban areas of Benin. *Malaria Journal*, **8**, 103.

Yamamura, K., Yokozawa, M., Nishimori, M., Ueda, Y., & Yokosuka, T. (2006). How to analyze long-term insect population dynamics under climate change: 50-year data of three insect pests in paddy fields. *Population Ecology*, **48**, 31-48.

Yukawa, J. (2000). Synchronization of gallers with host plant phenology. *Population Ecology*, **42**, 105-113.

Zainal, N. A., Zulkifli, N. W. M., Gulzar, M., & Masjuki, H. H. (2018). A review on the chemistry, production, and technological potential of bio-based lubricants. *Renewable and Sustainable Energy Reviews*, **82**, 80-102.

Zaller, J. G., Moser, D., Drapela, T., Schmöger, C., & Frank, T. (2009). Parasitism of stem weevils and pollen beetles in winter oilseed rape is differentially affected by crop management and landscape characteristics. *Biocontrol*, **54**, 505-514.

Zeileis, A., Hothorn, T. & Hornik, K. (2008a). Model-based recursive partitioning. *J. Comput. Graph. Stat*. **17**, 492–514.

Zeileis A., Kleiber C. & Jackman S. (2008b). Regression Models for Count Data in R. *Journal of Statistical Software*, **27**.

Zhou, W., Xu, C., Ge, T., McAuley, J., Xu, K., & Wei, F. (2020). BERT Loses Patience: Fast and Robust Inference with Early Exit. *arXiv preprint arXiv:2006.04152*.

Zhou, X., Perry, J. N., Woiwod, I. P., Harrington, R., Bale, J. S., & Clark, S. J. (1997). Detecting chaotic dynamics of insect populations from long-term survey data. *Ecological Entomology*, **22**, 231-241.

Zimmer, C. T., Bass, C., Williamson, M. S., Kaussmann, M., Wölfel, K., Gutbrod, O., & Nauen, R. (2014). Molecular and functional characterization of CYP6BQ23, a cytochrome P450 conferring resistance to pyrethroids in European populations of pollen beetle, *Meligethes aeneus*. *Insect Biochemistry and Molecular Biology*, **45**, 18-29.

Zimmer, C. T., Köhler, H., & Nauen, R. (2014a). Baseline susceptibility and insecticide resistance monitoring in European populations of *Meligethes aeneus* and *Ceutorhynchus assimilis* collected in winter oilseed rape. *Entomologia Experimentalis et Applicata*, **150**, 279-288.

Zimmer, C. T., Maiwald, F., Schorn, C., Bass, C., Ott, M. C., & Nauen, R. (2014b). A de novo transcriptome of European pollen beetle populations and its analysis, with special reference to insecticide action and resistance. *Insect Molecular Biology*, **23**, 511-526.

Zimmer, C. T., & Nauen, R. (2011a). Pyrethroid resistance and thiacloprid baseline susceptibility of European populations of *Meligethes aeneus* (Coleoptera: Nitidulidae) collected in winter oilseed rape. *Pest Management Science*, **67**, 599-608.

Zimmer, C. T., & Nauen, R. (2011b). Cytochrome P450 mediated pyrethroid resistance in European populations of *Meligethes aeneus* (Coleoptera: Nitidulidae). *Pesticide Biochemistry and Physiology*, **100**, 264-272.

Zlof, V. (2008). Recommendations and conclusions of the Ad hoc EPPO Workshop on insecticide resistance of *Meligethes* spp.(pollen beetle) on oilseed rape. *EPPO Bulletin*, **38**, 65-67.

Zuur, A. F., & Ieno, E. N. (2016). *Beginner's guide to zero-inflated models with R*. United Kingdom: Highland Statistics Limited.

Zuur, A., Ieno, E. N., Walker, N., Saveliev, A. A., & Smith, G. M. (2009). *Mixed effects models and extensions in ecology with R*. Springer Science & Business Media.