Spatio-temporal modelling of detection and spread of invasive plant pests

Amani Saleh Khalaf AL-Rubkhi. PhD Student

Mathematics and Statistics Department
University of Strathclyde, Glasgow

January 7, 2025

This thesis is the result of the author's original research. It has been composed by the author and has not been previously submitted for examination which has led to the award of a degree.

The copyright of this thesis belongs to the author under the terms of the United Kingdom Copyright Acts as qualified by University of Strathclyde Regulation 3.50. Due acknowledgement must always be made of the use of any material contained in, or derived from, this thesis.

Signed: Amani Saleh Khalf AL-Rubkhi

Date: Sep. 2024

Academic and professional engagement

1. Poster Presentation at International Conference:
    - European Conference on Mathematics applied to Biology and Medicine held September 19-23, 2022 in Heidelberg, Germany.
    - Conference on Innovations to Transform Drylands" held from 21-23 February 2023, at ICRISAT, Hyderabad, Telangana, India.
2. Presentation at International Conference:
    - Conference on Innovations to Transform Drylands" held from 21-23 February 2023, at ICRISAT, Hyderabad, Telangana, India.
3. Presentation to research partners
    - October 2022, I gave a talk titled "Detection of the 2018-19 Invasion by Fall Armyworm in India by Mobile Technologies: A Statistical Analysis." The audience included researchers from Progressive Environmental and Agricultural Technologies (PEAT) in Berlin, Germany GmbH, and the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT). The talk was delivered online.
    - August 2024 I presented an overview of my thesis to an audience from the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT).

Abstract

Invasive plant pests significantly threaten agriculture, ecosystems and economies. Effective control requires early pest detection and reliable spread assessment, which can be addressed through statistical analysis. This thesis develops novel statistical approaches to data analysis and modelling for two invasive insect species: Fall Armyworm (FAW) in India during the period 2018-19, and the Emerald Ash Borer (EAB) epidemic in the USA from 2002 to 2020.

The decision-making for prevention and control is often hampered by the lack of data or their low quality. Novel detection methods have been developed, but they have not yet been analysed rigorously. The FAW data are based on a citizen science approach utilising Plantix, an innovative method that integrates artificial intelligence (AI) with mobile technologies. Constructed by Progressive Environmental and Agricultural Technologies (PEAT) GmbH, it can provide comprehensive monitoring of geographical areas and early detection of pest invasions. However, there is no gold standard, and the data need to be statistically interpreted before they can be used to estimate prevalence.

For the EAB data, a different approach is needed as only an initial true positive case was provided from each observed infested county in the USA. However, we also have data on the host (ash trees) density and climate forcing. For the two cases, the main research objectives are: (i) developing a rigorous framework for estimating FAW prevalence and using it to estimate the true prevalence in different parts of India, and (ii) developing a continental-level model for the spread of EAB in the USA.

Both approaches apply frequentist and Bayesian techniques, using classification methods and several diagnostic performance tools to compare model outputs with data. A classification model, a bi-normal mixture, was used to estimate the True and False FAW observations, using the data classification by the Plantix mobile app and our assumptions. A Bayesian meta-analysis estimates pooled test sensitivity and specificity, assuming the logit sensitivity and specificity follow a multivariate normal distribution.

Four distinct methodologies were implemented to select the most appropriate model for estimating FAW prevalence, including frequentist methods and the Bayesian meta-analysis with stochastic sensitivity and specificity.

In the case of the EAB, a colonisation-dispersal model was adapted and utilised to include climatic (annual average of growing degree day), non-climatic (ash intensity habitat) conditions, and dispersal mechanisms. The model was fitted to the best available data, quantifying the uncertainty in the model and its predictions and assessing its performance in tracking the spread of EAB over two decades.

The thesis analysis yields key findings for both FAW and EAB. These findings classify positive and negative Plantix app observations as True or False, evaluate app accuracy and enable estimation of FAW prevalence. Additionally, the evaluation of the data sensitivity and specificity for each maize season is more accurate than for the entire period. Significant factors for EAB colonisation are ash species availability, and the adult EAB flights dispersal distance.

The results also highlight that the citizen science and mobile technologies can aid the government in early pest detection for effective spread control of FAW, EAB, and similar pests, and may even combined with inspector monitoring to limit the EAB spread.

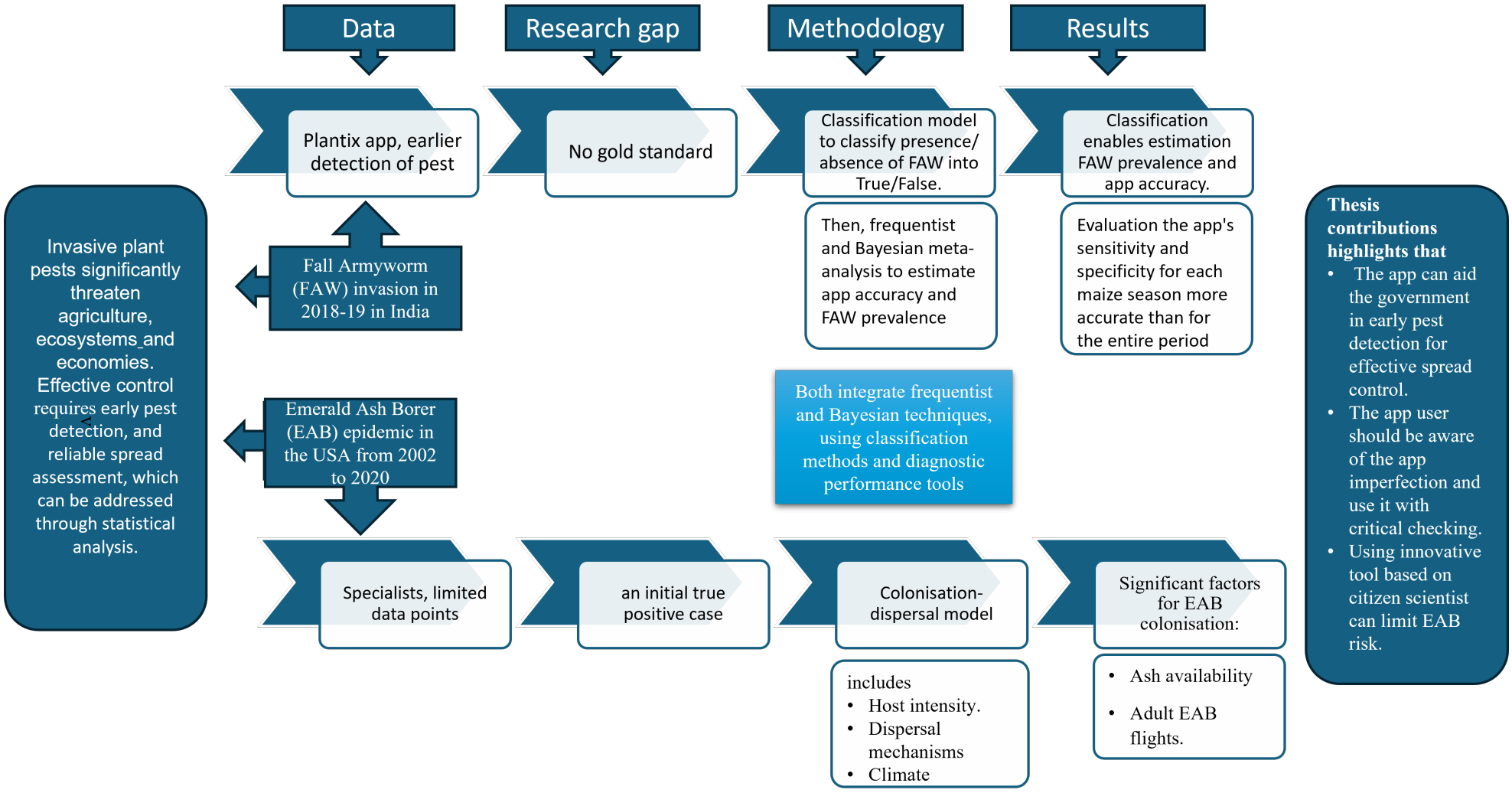


Figure 1: Graphical Abstract

Contents

List of Figures

List of Tables

Preface/Acknowledgements

I am deeply grateful to all those who have supported me throughout my journey to completing this PhD thesis. First and foremost, I would like to acknowledge my sponsor, Sultan Qaboos University in Oman. Without its support, this thesis would not have been possible. I would also like to sincerely thank my family for giving me the opportunity to travel away from home for years. Their support and encouragement have been the foundation of this journey, which has helped me to reach my dreams. I am deeply saddened that my uncle, who supported me throughout my entire PhD journey and took pride in my achievements, will not be with me for my submission and graduation, as he passed away on August 8, 2024. I want to take this moment to express my heartfelt gratitude for everything he has done for me, and I wish to remember this feeling as a way to honor his support.

I would also like to acknowledge my supervisor, Adam Kleczkowski. His discussions and support were instrumental in helping me complete this thesis. I would appreciate thanking Dr. Srikanth Rupavatharam, Head of Innovations Hub & Senior Scientist, ICRISAT, India, for his academic support and for uplifting my spirits. He always makes me feel better.

Chapter 1

Introduction

1.1 What are invasive species?

Invasive non-indigenous (non-native, naturalised, exotic, alien) species are species that have moved accidentally and introduced into a new geographic area. They also can persist, reproduce, spread and cause negative impacts in the local ecosystem and biodiversity [6, 7]. They can disrupt the ecological balance when they encounter mismatched abiotic and biotic factors in their native habitats [8]. The IUCN (International Union for Conservation of Nature) Red List of Threatened Species and the 2019 IPBES (Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services) Global Assessment Report on Biodiversity and Ecosystem Services highlight that the invasive non-native species are one of the primary reasons that leads to a reduction in biodiversity and causes species extinctions. In addition, they rapidly threaten food and livelihood security [9]. Additionally, Early et al. (2016) [10] found that in the 21st century, these species are highly susceptible to invasive on the one-sixth of the global land surface. Hence, they are a global concern [11].

In addition, the alien species can be transported by abiotic dispersal such as wind, animal, or water. They can also move independently [7] or due to human actions such as by carrying non-native organisms in material behind native area [7, 12]. It is possible for airplane wheels to carry species to a new region. Additionally, cargo ships can transport marine alien species into new areas, since they sometimes carry ballast water to stabilise the ships load. Then, the ballast water that may contain marine organisms is released into the port. Moreover, cargo ships and trucks or cargo packaging materials could carry non-native species into containers. About 51.8 percent of solid wood packing materials

shipments had alarming infection rates where can carry wood-boring insects, according to a collaborative report published by the United States Department of Agriculture (USDA), the Animal and Plant Health Inspection Service (APHIS), and the United States Forest Service (USFS) [12]. Further, the international trade of live plants has been one of the main factors to introduce alien forest insects and pathogens into the USA and Europe. For example, between 1860 and 2006, around 69 percent of non-native forest insects and pathogens entered to the USA through international trade of live plants [13].

1.2 Threat of invasive insects

Although invasive non-indigenous species can be plants, insects, vertebrates, or marine organisms, this work will focus on insects. Non-native insects have increased rapidly around the world, which endanger native biodiversity, ecological, economic, and human life [14]. Further, invasive non-native insects can, through direct interactions, disrupt the native ecosystem balance and contribute to decrease in native biodiversity. Moreover, they consume plants or spread disease [15, 16]. In addition, they might act as parasitoids and lay their eggs inside other insects. After hatching, the emerging non-native larvae harm the native ones [16].

In addition, non-indigenous invasive insects have had marked economic effects on different sectors, such as forestry, agriculture, society and trade. They have been estimated to cost the world economy more than US$ 70 billion yearly [17, 18]. For exapmle. in agriculture, they incur billions of dollars of losses by reducing yields, increasing pesticides usage, spreading plant pathogens and imposing trade restrictions [19]. United States of America (USA) governments spent billions of dollars per year to eliminate non-native forest insects, to limit their distribution to new areas, and to eradicate at risk and dead trees from public areas for public safety [13].

Moreover, householders incur financial costs to remove or replace infested trees. Their property prices also may be reduced due to the risk of the infested trees and less attractive home views [13]. Further, global health consumed around US$ 6.9 billion per year to treat human disease because of invasive insects [20]. In spite of the above given costs, the cost of non-indigenous invasive insects is still underestimated. Difficulty in quantifying costs, insufficient government funding and lack of international coordination or cooperation are factors that cause this underestimation of costs of the alien invasive

insects [20].

In this thesis, the focus is on two alien insects that have recently become serious pests in some parts of the world: the Fall Armyworm (FAW, *Spodoptera frugiperda*) on maize crop in India, and the emerald ash borer (EAB, *Agrilus planipennis*) in ash species in the USA.

1.2.1 Fall Armyworm, crop non-native insect, in maize

The Fall Armyworm (J. E. Smith), is a highly mobile and dangerous pest, which was first recognised by Sir James Edward Smith [21, 22]. The pest is classified from the genus *Spodoptera* known as armyworms, and the Noctuidae group which is one of the causes of agricultural financial losses around the world [21]. It infests a wide range of host plants with approximately 100 recorded types under 27 families. However, the preferred species is the *Gramineae* family of plants, including the economically essential ones such as maize, millet, sorghum, sugarcane, rice and wheat. Different reports indicate that there are other crops such as cowpea, groundnut, potato, soybean and cotton, which are infested by this pest [23]. It is notable that between all these host plants, FAW caused the most damage in maize [24]. Hence, the subsequent sections will discuss the FAW around the world and particularly in India, and at the end will highlight the significance of maize.

FAW around the World

The Fall Armyworm is native to the tropical - subtropical regions of the American continent, where it is found in Mexico, Brazil, the USA and Argentina [25, 26]. In Brazil, the yield of maize crops has been reduced by 98–100 percent due to Fall Armyworm [27]. In addition, in 1845 in western Florida, FAW caused massive damages to various crops such as corn, sugar cane, and rice [28]. One farmer in 1870 consumed US$1,000 because of damage caused by the FAW to several crops [28]. Hence, it posed a considerable damage even in its original habitat.

In 2016, it was found in West and Central Africa and subsequently invaded all the countries of sub-Saharan Africa [25, 26]. In addition, in 2018, it moved to Asia through India and then expanded to Bangladesh (December 2018), Myanmar (December 2018), Nepal, Sri Lanka and China (January 2019) and South Korea and Japan (July 2019) [29]. The map in Figure 1.1 shows FAW distribution around the world since 2016. This

pest, which only grows up to 2-cm long, extended to more than 50 countries in Africa and Asia damaging crops, especially maize [30, 29, 26]. In February 2020, FAW also invaded Australia and attacked more than 350 commercial and non-commercial hosts [31].

In fact, the factors that helped FAW to spread fast over the continent was the importation of non-genuine species for economic purposes. Moreover, unintentional infected plants arrived in shipping containers were carried by tourists or hidden in the imported ornamental plant soil [21]. Other factors that contributed to its spread were its predilection to attack many crops, especially the maize. Furthermore, it has ability to produce numerous eggs, and its capability to travel over wide distances [21].

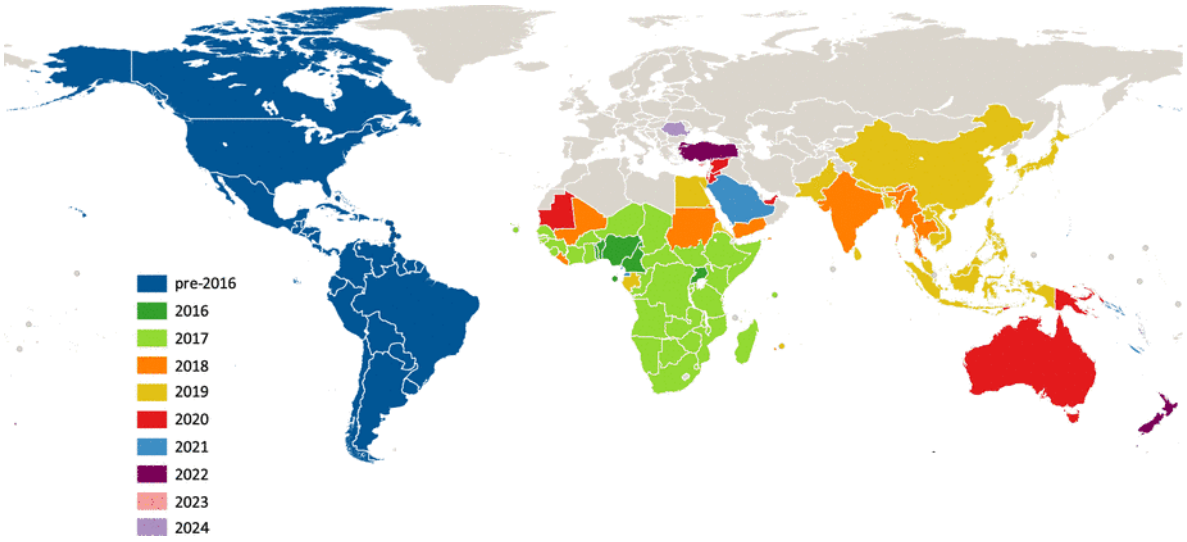


Figure 1.1: FAW Invasion around the World since 2016, taken from [1].

Fall Armyworm invasion of India

In mid (May-June) 2018, *Spodoptera frugiperda* was detected first on the maize crop in many locations in Karantaka, India. Then, the pest expanded to all southern states, then to western Maharashtra and Gujarat, and eventually to eastern states [30]. Between July and August 2018, severe damage was reported in Chikkaballapur, Hassan, Shivamogga, Davanagere and Chitradurga which are located in Karantaka [23]. The pest was also observed in Andhra Pradesh, affecting the maize growing areas of East and West Godavari districts, Srikakulam and Vizianagaram in August, 2018 [30]. In just two years, FAW has grown in most parts of India [32].

Life cycle of the Fall Armyworm and visible characteristics

The complete worm lifecycle of FAW varies and depends on the prevailing natural conditions. It averages 30 days in summer, 60 days during the spring (Rabi) and autumn (Kharif), and during the cold season (winter), it can be prolonged from 80 to 90 days [33]. Male and female adult moths mate nocturnally; from dusk to midnight, and are most active during warm, humid evenings. The adult female moth lays between 1,500 and 2,000 eggs in her life time, on the inner side of the whorl or on the under surface of the leaf [33, 23]. She deposits her eggs in a mass which is arranged in a layer, or sometimes layers, although most eggs are distributed on a single layer. The number of eggs per mass is very different but is usually between 100 to 200. During laying, the female arranges fine protective grayish scales which are loosened from her body to cover the eggs and the egg mass, over which this cover has a moldy appearance [33, 28]. The eggs are brownish yellow and are of dome shape, where the diameter is about 0.4 mm and the dome height is 0.3 mm [33]. The duration of the egg stage ranges between two and ten days, where it depends on temperature conditions, rather than humidity, whether in a dry or moist place [28].

After hatching, the larvae develop through six instars; around 14 days in the summer with 30 days during cool weather. The instars differ slightly in physical appearance and pattern. The average development time of each instar in order was determined to be 3.3, 1.7, 1.5, 1.5, 2.0, and 3.7 days at 25°C [33]. The larvae in the early instar (1st instar) are greenish with a black head, where the head width is around 0.35 mm and its body length is about 1.7 mm. Larvae in the second and third instar are orangish and have a black head of 0.45 to 0.75 mm in width and attain a length of about 3.5 and 6.4 mm, respectively. In the second, and mainly the third instar, the dorsal surface of the body becames brownish, then white lines begin to form. Larvea from the fourth to the sixth instars have head widths of 1.3, 2.0, and 2.6 mm, respectively, and their heads are mottled reddish brown and white. Moreover, the body lengths are 10.0, 17.2, and 34.2 mm, respectively. The body colour is brownish with subdorsal and lateral white lines. Some spots appear on the dorsal, which are darker than the body colour and have spines. The faces of the 6th instar larvae (mature larvae) have unique mark on Y form, and the epidermis of the larva is rough [33].

After leaving the plants, the larvae pupate in the soil to begin a pupal stage. The duration of a pupal stage ranges eight to nine days in summer and 20 to 30 days in cool

weather. The pupation stage happens in the soil at a depth range of two to eight cm. The larva forms a loose cocoon by tying together particles of soil with silk. However, if the soil is hard, the larva pupates on the soil surface by mixing leaf debris and other materials to construct the cocoon. The cocoon has an oval shape and is 20 to 30 mm tall. Gradually, the adult moth emerges to start to the adulthood stage and a new life cycle. Moths have a wingspan of 32 to 40 mm. The forewing of a male moth is shaded brown and grey, with triangular white spots at the rim and close to the centre of the wing. However, the forewing of the female moth is between a uniform greyish brown to a fine mottling of grey and brown. The back wing in both sexes is iridescent silver-white with tight dark edges. The female deposits most eggs in the first four to five days of her life, but some oviposition occurs for up to three weeks. The duration of an adult's life on average is approximately ten days, with a range of 7 to 12 days [25, 33, 23].

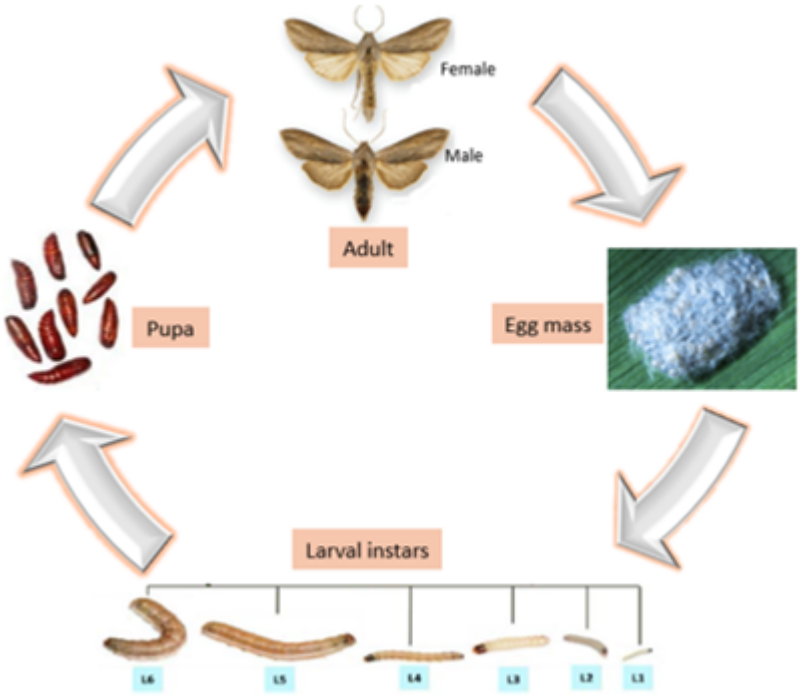


Figure 1.2: Life cycle of Fall Armyworm (FAW)

Damage and management

FAW can damage maize during all growth stages, whereas it causes a greater threat in the vegetative period or young crops, and FAW is more destructive in larvae stage since it may also attack tassel and developing ears. Larvae in all instars cause damage by consuming foliage. The larvae in the early instar (1st instar) feed on the eggshells firstly

and then invade chlorophyll (green tissue) on upper leaves to form a silvery transparent membrane. In the second and third instar, larvae start to form window pane (hole) on the leaves and leave moisture resembling sawdust in the funnel and upper leaves, which they start eating from the edge, working their way to the inside. Larvae in the third instar to the sixth instar cause heavy defoliation and leave ribs and stalks on the plant, as well as a large amount of faecal matter. Some larvae can also get onto neighbouring plants [25, 33, 23]. To sum up, as larvae grow, their feeding rate increases, the hole sizes and the amount of facal matter increase. Larvae in the first to the third instar consume 2 percent of the total foliage since they are pretty small, while for the fourth, fifth and sixth instars, they eat 4.7 percent, 16.3 percent and 77.2 percent, thus heavily defoliating the crop [29, 24].

Substantial economic losses in agricultural biodiversity, human and animal health are caused by FAW [23, 30]. In Latin America, FAW caused up to 73 percent of yield losses in maize crops [29]. Further, over 44 countries in Africa and just in a two-year period (2016-2017), the damage of FAW in maize led to a financial loss exceeding US$2–5.5 billion [27]. 13.5 million tons of maize, valued at $3 billion, was the estimated impact of FAW in sub-Saharan Africa during 2017-2018. It is greater than 20 percent of its maize production [32]. Within two years, in Ethiopia FAW contributed to the loss of maize production by an amount 0.67 million tonnes. In the absence of this lost maize, four million food insecure households would have been able to meet their maize consumption needs [34].

In India, there has been widespread concern about FAW in maize fields since mid 2018. It spread to more than 90 percent of maize paddies within 16 months [35, 29]. Worth mentioning is the fact that for every 5 to 10 percent drops in production, India loses 37,000-75,000 tonnes of maize [29]. In Kharif 2018, 17,394 ha out of 462,322 ha of the actual sown area of maize was infested by FAW and 22,072 hectares (ha) out of 78,982 ha of the actual sown area during Rabi 2018-19 [30]. In addition, Suby et al. [29], mentioned that in 2019 Karnataka recorded the largest infested area with FAW (211,300 ha), followed by Telangana (24,288 ha) and Maharashtra (5144 ha). Furthermore, FAW caused economic damage in the rainy and post-rainy seasons of 2018 and 2019 in Andhra Pradesh, Karnataka, Maharashtra, Tamil Nadu and Telangana. However, FAW did not cause economic damage in fields with temperatures less than 10°C or more than 40°C. As a result, there was only a minor FAW infestation in northern Rajasthan, Haryana

and Punjab [29]. India is the world's seventh-largest maize producer, which it typically exports to Asia. As FAW infestation area increases over time, the loss will shift India to the import of maize [30].

Managing the pest is difficult since the worm can be seen in all stages at the same time because of the continual generations [30]. However, there are ways to minimise this invasive insect pest such as using biological control agents or their natural enemies which are Telenomous sp. and Trichogramma sp, as well as quarantine restictions, all of which could result in reducing the pest globally [23, 30].

FAW weather conditions

FAW can live throughout the year, diversify its diet, and survive in difficult and harsh conditions by emigrating to various areas or hiding to return when the conditions are appropriate. Despite this dangerous threat, the moth needs suitable weather conditions to survive. Warm and humid conditions help its reproductive capabilities, while extreme temperatures or excess rainfall cannot be tolerated. Kenya, for example, is not affected by the pest due to the heavy rains experienced in March 2018 [21]. The moth usually attacks crops during the larval stage, but its lifecycle seemed to be broken because of the rains. Since the pest behaves differently from one area to another and from one season to another, scientists should study these behaviours in diverse environments to come up with a prediction for each season [30].

Significance of maize

Maize is an essential cereal in India as well as in many countries in the world. From 2018 to 2021, statistics indicate that maize is sown in 165-170 countries in areas of about 180.63-188 mha (Million hectare), with the annual production ranges between 1060 to 1134 mt (million tonnes) [29, 36]. In 2021, China has the largest maize area followed by the USA, where both cover 39% of the world maize area. However, the USA is the highest maize production followed by China, where they are contributing 34 percent and 22 percent of world maize production [36]. Worldwide, maize is used as food, feed, fodder and raw material. Moreover, the main advantage of maize is that it can be grown in a moderate climate and planted from sea level elevation to 3000 m above sea level [36].

In India, maize is the third staple crop after rice and wheat, and it covers 4% of

the maize world area [37, 36]. In addition, India has been among the top ten maize production in the World since 1961. Since 2005, India has ranked 4th area under maize with 9.2 mha of land with a production of 28 mt. Yet, in 2021, India ranks 7th, while the productivity is above 3 t/ha [36]. Further, Indian maize contribution is about 9 percent of local production and around 2 percent of global production [37], where around 50 to 60 percent of local production is used as food for people and feed for cattle. Moreover, about 30-35 percent of the production is consumed for poultry, piggery and fish meal. Additionally, 10 to 12 percent of it is used in wet milling industry such as starch and oil, and around 3 percent in dry milling such as corn bread and corn chips [38].

Maize was a rainy season or Kharif crop in northern India before 1980 (1950-1979), and it was grown mostly in the states of Uttar Pradesh, Bihar, Rajasthan and Madhya Pradesh [36, 24]. After 1980s, Rabi (Winter) maize has become important in coastal Andhra Pradesh, Bihar, Telangana, West Bengal and others [36]. Simultaneously, it was noticed that there is a considerable shift in area towards peninsular region which represents now of about 40 percent of the total area under maize and over 52% of production [36, 24]. Since 2017-18, the major maize growing states represent of about 80% of the total maize area of the country which are Karnataka (14.8 percent, 1.22 mh, 3.31 mt/y), Maharashtra (10.9 percent), Madhya Pradesh (10.8 percent), undivided Andhra Pradesh (10.4 percent), Rajasthan (10.6 percent), Uttar Pradesh (8.3 percent), Bihar (7.9 percent), Gujarat (5.0 percent) and Tamil Nadu (3.6 percent). However, in many of these states such as in Rajasthan (1.6 t/ha) and Gujarat (1.6 t/ha), maize productivity is quite low, while that in Uttar Pradesh (1.7 t/ha), Madhya Pradesh (1.9 t/ha) and Maharashtra (2.3 t/ha) is under the national average of 2.6 t/ha [24].

Moreover, maize can be grown in a mild climate [39]. It can also be grown well in semi-arid, humid, hot dry or hot moist conditions. In addition, it can be planted in all type of soils, and the best range of soil pH is 7.5 to 8.5 [40]. Therefore, production of maize in India is round-the-year [36], and maize can be grown in all seasons in most maize fields. It can be grown in Kharif (monsoon), post monsoon, Rabi (winter), summer and spring [39], while Kharif season is the most suitable period to sow maize [24]. On the other hand, farmers use irrigation during Rabi and spring seasons to achieve higher yield [39]. However, in Bihar, West Bengal and Peninsular India, maize often grows in Rabi season, while in Punjab, Haryana and western Uttar Pradesh it grows in Summer season, requiring more water. Therefore, 80 percent of maize is Kharif maize,

19 percent is Rabi maize and 1-2 percent is Summer maize. In spite of that, Rabi maize is growing faster than Kharif maize and with double yield [24, 40].

It is worth highlighting that in 1950-51, maize production was about 1.73 mt, and increased to reach 28.75 mt in 2017-18. This rise is because maize area has increased 2.97 times, yield 5.6 times and production 16.64 times compared with the beginning of the period in 1950-51. India aspires and plans to reach its production of 50-60 mt by 2025. However, climate change is one of the challenges that cause stresses and fears to achieve this goal. Along with that, from May 2018 Fall Armyworm has threatened maize crop [24].

1.2.2 Emerold ash borer, forest non-native insect, in ash trees

Emerold ash borer (EAB) is a wood borer beetle that feeds on ash trees (*Fraxinus*) [41]. It is indigenous to East Asia (north east China, Japan, Taiwan, Korea, Mongolia and the Russian Far East) and was discovered in the southeastern Michigan, USA in 2002 [42, 43, 44, 45, 46]. However, it is suspected to have arrived ten years prior with solid wood packaging material from Asia and was only identified when ash trees began to die in significant numbers between 2001 and 2002 [42, 44]. This is because *A. planipennis* completes its life cycle inside ash [41], and is extremely difficult to detect and can remain unnoticed for several years after the infestation [44]. At the end of 2002, it was obvious that between 5 and 7 million ash trees were dying, declining, or dead in six counties of southeastern Michigan [42]. Over the next five years, approximately 20 million trees had been killed by EAB in Michigan state alone [47]. These alarming figures indicate that the EAB has had a substantial ecological and economic impact in the infested area [48]. Following is a review of the literature on EAB biology, impacts, and management options.

Biology: life cycle of the EAB and visible characteristics

EAB life cycle involves four distinct stages: egg, larva, pupa, and adult. The development time of the EAB life cycle is flexible, which allows *A. planipennis* to establish in a variety of climates and other environment factors [49, 50]. For example, a weakened tree (e.g. by girdling) has a shorter generation time than a healthy tree [50]. The life cycle of the EAB typically lasts one year, while it occasionally lasts two years. One of the contributing factors is that the 2-year development occurs occasionally when oviposition

takes place in the latter part of the summer and the larvae do not reach the prepupal stage before winter [42, 51]. In summary, the EAB's life cycle is one or two years in North America [49].

From infested ash trees, adults emerge ranging in length from 8.5 to 12.5 mm with bright green coloration [52]. Prior to mating, adults feed on ash foliage for at least a week to reach sexual maturity [53, 50]. Following mating, female adults bore into the bark to lay their eggs, producing between 40 and 70 eggs at once [49]. Individual eggs are laid within cracks and crevices in the bark, or beneath bark flakes, and hatch in approximately two weeks [49]. Each egg is about 1 mm in diameter, and its color develops over time from white to amber [54].

Once the eggs hatch, larvae appear in the bark [54]. A larva will undergo four instars (stages) over time, where depelovement periods depend on the temperature and other factors in the environment [49]. The four-instar larva can be distinguished based on their sclerotized structure [54]. They tunnel to make their way to the cambium where they feed, etching a serpentine gallery in the phloem and outer sapwood [54, 52]. Through tunneling, they slowly cut off moisture and nutrients to the higher parts of the tree [52].

Larvae of the fourth instar chew pupation chambers in the outer sapwood or bark during late summer or autumn and fold their bodies into J-shaped larvae, the stage at which they overwinter [54, 53] in an indefinite diapause as prepupae [53]. However, if the J-shaped larva does not reach the appropriate size and development stage by late fall, it will delay pupation until the next summer [54]. In the spring, the prepupae will develop into pupae. A pupa gradually develops into an adult over the course of about one month while still in the pupation chamber. When adults emerge from the tree bark, they appear from D-shaped exit holes (2–3 mm in diameter) [54, 49] and are capable of immediate flight. The adults consume ash foliage during their lifetime and are most visible during hot afternoons (3–6 p.m.), flying around ash tree trunks and landing to reproduce [54].

EAB adult flight and natural influences

EAB can migrate from one area to another either at the adult stage of its life cycle or via anthropogenic movement of ash materials [42, 55]. According to a laboratory study for assessing EAB flight ability using computer-monitored flight mills, the ability of

EAB to fly differs between fed and unfed, as well as mated and unmated of both sexes. A male EAB flies more frequently and farther than an unmated female. Mated females, however, are likely to increase a population's range through dispersal and reproduction. According to the flight mill studies, mated females which were allowed to feed on ash leaves between flight periods, flew further and longer per day than unmated females (almost 2.5 times farther) or males. The average flight distance for mated females was 1.3 kilometers per day. The average flight distance of fed, six-day-old female beetles ranged from 71 m to 2.3 km, compared to 53 m to 5.2 km for males. The results show that mated females are likely crucial in the spread of EAB, especially if flight mill recordings are underestimated [55]. The results, however, cannot be applied directly to field scenario, because environmental factors can also impact adult flight behaviour [42, 55].

Environmental factors such as ash tree distribution, density, and condition can affect EAB adult flight and dispersal [42, 55]. For example, adults often prefer stressed trees and trees expose to full or nearly full sunlight. Additionally, an indirect factor is a host's phloem quality and availability which impact larval development. Subsequently it influences adult flight [55]. Additional elements are weather conditions such as wind or geographical features including mountains or sea [42].

Destructive impacts of EAB

The emerald ash borer is considered to be one of the most destructive forest insects ever to invade the USA [14]. It has caused major damage in ash species in the USA since 2002 [49]. It infests various ash species, though the impact differs substantially between individual species [14]. *A. planipennis* poses a threat to all of North America's ash species, including at least 16 native varieties, as well as naturalised species and cultivars used in landscaping [42]. The larvae have a considerable impact on ash trees because they disrupt the tree's ability to transport water and nutrients. As a result, this leads to canopy dieback, bark splitting, and ultimately tree death [49]. Dieback in affected stands usually occurs within six years, and 50% of it occurs within four years [56]. As a result, the damage caused by EAB can have obvious ecological, aesthetic, and economic impacts.

The effect of EAB on ash species have caused both direct and indirect ecological impacts in the USA [42]. When ash trees die, the composition and structure of the forest

change [54]. This affects animal species that depend on ash for nesting sites, food, or shelter. Ash trees are important components of many forest ecosystems, providing habitat and food for a variety of wildlife species [49]. At least 43 monophagous species, native to North America, that may be at risk of coextirpation as ash is eliminated from the ecosystem [14]. Moreover, ash trees provide opprturnity to browse, thermal cover, and protection for a range of wildlife [42]. Furthermore, water quantity in the soil may change, due to ash death [48].

Additionally, dead and dying ash trees create both economic and safety risks to people and property [54]. Therefore, EAB spread attracts the attention of not only entomologists, ecologists, and forestry experts, yet also the general public [46]. The EAB spread has resulted in hundreds of millions of dollars being lost by municipalities, property owners, nursery operators, and the forest products industry [57]. Dying or dead trees drops the value of house, because aesthetic value reduces [48], This is because ash trees provide thermal cover and an attractive view of a home. Furthermore, there is the risk of falling ash trees on humans and houses. Consequently, both governments and homeowners need to remove drying or dying ash trees. In addition, homeowners suffer a loss in the value of of the timber on their land, which is significantly lower than that of healthy trees [48].

EAB presents a major economic concern as it causes direct costs arising from the removal and replacement of dead or dying ash trees and other management techniques [48]. For example, Kovacs et al. (2010) estimates that the cost of removing or treating less than half of the infested ash in the USA cities will be more than $10.6 billion by 2019 [58]. Moreover, removing ash trees was estimated to cost between $20 and $60 billion, without taking into account the costs associated with their replacement. Additionally, ash trees account for lumber industry in the eastern USA, with an estimated stumpage (standing timber) value of $25 billion [42]. This underscores the negative economic impacts of the EAB, if it infestes ash trees. It is worth noting that there have also been indirect costs arising from the loss of political considerations in the allocation of government funding for surveys, research, and outreach activities [48].

To conclude, EAB poses a significant threat to ash tree populations in the USA, by causing ecological, aesthetic, and economic damages. However, a number of effective management strategies are available to mitigate the impact of this invasive species that will be explored in the following paragraphs.

Management strategies

A number of strategies have been developed to manage EAB infestations in the USA, including quarantine, mechanical control, biological control, cultural practices, tree removal and replacement, and chemical treatment.

Quarantine

One of the first management actions taken after the detection of EAB in the USA in 2002 was the creation of quarantine zones. In areas where EAB is present, quarantine measures had been implemented to limit the movement of potentially infested materials such as firewood, nursery stock, and cut logs. A quarantine program prevents the spread of disease across long distances and slows down the rate at which new infestations develop. It may be effective when combined with other management strategies such as early detection and the removal and destruction of infested trees [14]

Mechanical control

Ash tree removal is an effective approach in controlling EAB infestations [59, 14, 60]. Moreover, this practice is often necessary when trees are heavily infested and threaten human or property safety [14, 60]. Further, it is an effective method when EAB is not yet widespread [60]. In other words, implementing this strategy may not be appropriate in outbreaks in which the EAB has already established a strong presence [60].

Removing ash trees can slow spread of EAB and provide protection to the surrounding healthy ash trees in the area. As ash trees are removed, there would be fewer habitats for EAB adults to breed and larvae to feed and survive. As a result, this would reduce the EAB population and hinder its spread [55]. However, eliminating ash trees can be expensive and have negative effects on the environment and community aesthetics, resulting in higher heating and cooling costs, lower property values, and changing wildlife habitats [60]. To mitigate these impacts, replanting with other species can promote ecological function and reduce future infestation risks [59, 60]. Additionally, if eradication efforts fail, conventional biological control methods will be needed to suppress EAB populations [61].

Biological control

Biological control technique, which relays mainly on natural enemies, can be effective in reducing EAB growth in population and spread [49]. There are predators, pathogens or parasitoids as effective natural enemies of EAB [42]. Clerid beetles are an example of predators and have been observed attacking EAB larvae [42].

The parasitoids are affecting either eggs or larvae [61]. Encyrtid parasitoids have been extracted from *A. planipennis* eggs, making them potential natural enemies of the EAB [42]. *Tetrastichus planipennisi* and *Spathius agrili* are two examples of larval parasitoids [61]. These parasitoids were first discovered in China, where EAB is native, and were later introduced to North America as part of a biological control program to help manage EAB populations. Both have been shown to be effective at reducing EAB populations in laboratory experiments and field trials [61].

Additionally, early field surveys held in the state of Michigan, USA and other newly infested sites, found that native North American parasitoids attacking EAB were marginally effective, resulting in no more than 5% parasitism being observed [62]. According to Cappaert (2009), *Atanycolus cappaerti* Marsh and Strazanac (Hymenoptera: Braconidae), a newly described, native North American parasitoid, parasitizes the EAB at two sites near Fenton, Michigan and had parasitism of up to 71% between 2007 and 2008 [63]. During 2009 and 2010, Roscoe (2016) estimated that Phasgonophora sulcata Westwood (Hymenoptera: Chalcididae) caused up to 35% persistence at some heavily infested sites in two cities in Canada [64]. Therefore, while biological control has potential as a strategy against the emerald ash borer, more research is needed to determine its feasibility and effectiveness in practice.

Chemical control

Controlling EAB infestations by chemical means is an important approach to managing infestations and ensures the survival of ash trees in an area [60]. The application of insecticides is one of the most commonly used chemical control methods. A variety of insecticides have been used to effectively control EAB populations, including systemic insecticides, such as emamectin benzoate and azadirachtin, and neonicotinoids, such as imidacloprid formulations and dinotefuran [60]. These insecticides are typically applied to the trunks or soil around the base of ash trees [65, 66, 60]. They are absorbed by the tree and delivered to the foliage, where they are consumed by feeding EAB larvae or

adults, leading to their death [65, 66]. The SLAM Pilot Project was undertaken between 2009 and 2011 in Michigan, and it concluded that emamectin benzoate can be highly effective at suppressing emerald ash borer infestations and slowing EAB population growth, also slowing the process of ash decline and mortality [66]. Factors such as the size of the tree, the timing of insecticide application, and the development of resistance may limit the effectiveness of insecticides [49].

Despite chemical control being an option for managing EAB populations, it is not widely used due to its high cost and environmental impact. Researchers are instead focusing on the introduction of the natural enemies of EAB as a more sustainable and effective means of reducing EAB populations in the forest ecosystems of North America [62]. Multiple management strategies are necessary for effective EAB management, minimising insecticide reliance and promoting sustainable EAB management [49].

1.3 The role of spatio-temporal data in invasive insects

Due to the non-native insects posing a significant threat, it is important to study them statistically. Data collection and analysis are essential for this task. In particular, spatial-temporal data is important, because it offers valuable understandings into patterns and variations species distribution over time. Consequently, the analysis can be helpful in early pest detection and selecting pest control regulations. It can also provide reliable spread assessment and identify priority for site-specific pest management [67]. Therefore, it is invaluable for researchers and policymakers to assist in mitigation of economic and ecological impacts.

1.3.1 Advantages and disadvantages of data collection by human experts

Precise pest insect identification tools are essential over time and space. Traditionally, pest identification methods rely on expert visual inspection, and field visits to collect essential information. Subsequently, it often involves expertise of entomologists. Commonly, the collection data by experts record an accurate invasive species type. Yet, because of the wide number of insect species, skilled entomologists may find it challenging in pest identification. Moreover, visual inspection by experts is not always practical or it may not be cost-effective or ethical [68, 69]. In addition, the accurate time and the

full extent of the observations in a specific location may not be possible to be recorded by experts. This is because the availability of the experts in the species location may be delayed due to a lack of experts or their late realisation of the species' existence [70]. Additionally, the global shortage of entomologists makes timely and accurate pest identification difficult. In particular, insufficient number of entomologists can be an increasing problem in developing countries with vast agricultural enviroments or remote locations such as forests.

1.3.2 Citizen scientists (nonprofessional scientists)

In response to the human experts data challenges, recent complementary approaches commonly involve citizen scientists supported by artificial intelligence (AI) smartphone applications [71]. Citizen science (CS) is data collection and scientific thinking by a volunteer in the biodiversity and environmental fields who monitors and collects ecological observations and physical specimens to expand the knowledge and database, supporting scientific researchers in their research [72]. Therefore, citizen scientists can minimise the time of data collecation and increase the accurate estimate of the occurrence time [71]. Additionally, they can collect large quantities of data more quickly, although this data might be less accurate compared to data collected by human experts.

Citizen scientists (nonprofessional scientists) are expected to fulfill some criteria. As a minimum, they should have interest, skills, and enthusiasm for the project's goals to ensure a successful project. Moreover, providing opportunities for citizen scientists as training and mentoring skills to develop their expertise is crucial, especially for new projects [72]. Additionally, it is essential for citizen scientists in the agricultural field to be under the supervision of scientists and experts to collect and share real and accurate data to minimise the risk of plant diseases [73]. Citizen scientists can use their smartphones to take photos of the plant symptoms. Then the image is sent via the internet to a human expert for diagnosis.

Alternatively, smartphone applications can be based on artificial intelligence and machine learning methods that enable autodetection and identification of pests and diseases [74]. In particular, the plant mobile applications either enhance scientific research or provide knowledge and information to farmers. The first one was created to provide information to farmers and for farmers to provide information to researchers. These type of apps serve users by supplying most or all of the following advantages:

standardizing data collection, prompting users through additional useful data collection questions, adding the ability to take photos associated with the record, automating Global Positioning System (GPS) reporting, and aggregating all data in a single, easy to use online interface. Purdue Plant Doctor (Purdue University, 2016) and Plantix (PEAT, 2016) are examples of this type of application. The second available type of plant-related mobile app does not contribute to ongoing scientific research. Instead, the apps serve as identification aides and/or act as repositories of educational information already publicly available, such as Forest Tree Identification (Discovery Green Lab, 2019), PictureThis - Plant Identifier (Glority LLC, 2019), About My Woods (Innovative Natural Resource Solutions LLC, 2019), and SEEDN (Bugwood, 2019). These types of apps are useful, but none of it focuses on facilitating scientifically meaningful collaborations between non-professional and professional researchers [75, 76].

Further, in 2018, the highest number of agricultural apps were used in the USA, Brazil and India. The large number of apps in the USA and Brazil was likely due to the strength of the agricultural sector, country size, and popularity of mobile devices. All these factors spurred public and private institutions to create applications for solving agricultural problems. In India, apps were introduced in 2015 and were free, but only accounted for one percent of total apps in 2018 [76].

Furthermore, the integration between citizen scientists and scientific research applications, especially plant diagnosis mobile applications, helps to monitor the health of the crop, to analyse soil, to suggest appropriate pesticides at right time before large scale incidence of disease, and to determine the optimal time for both planting and harvesting periods. Farmers can find answers to all their queries and receive relevant advice and recommendations to their specific farm related problems. These smart practices lead to the best yield and increase the accuracy of the prediction of yield. Agrio, Plantwise, Smart Scout, Veg Pest ID, Purdue Plant Doctor and Plantix apps are examples of plant disease diagnosis applications [77, 78, 68].

The Purdue Plant Doctor helps diagnose a disease, and is based on a questionnaire and decision trees in order to arrive at that diagnosis [71]. Leaf Doctor is a system for performing quantitative assessments of plant diseases. In addition, Pestoz (Creotix, India) diagnoses diseases from images of vegetable and crop plants with a primary focus on India. Furthermore, Plant Village Nuru is a system for diagnosing viral diseases of cassava and damage caused by FAWs on maize [74].

1.3.3 Plantix app

Plantix (PEAT GmbH, Germany) is a system for diagnosis of diseases, pest damage and nutrient deficiencies on crops. Plantix app is a free Android application, and it is an expert system trained to recognize a large number of diseases, to identify pests and pathogens, to define crop type, and nutrient deficiencies in soil. It communicates with a remote database (server) for higher accuracy where can it detects 500 types of plant damages. It also provides weather forecasting for the next five days in the current location of the user. Plantix app was created by a tech startup, Progressive Environmental and Agricultural Technologies (PEAT) in Berlin, Germany in 2015. The PEAT target is to use technology to support cultivators worldwide, explicitly in developing countries, to promote their agricultural production by diagnosing and treating disease [79, 80, 81, 82].

Accordingly, the Plantix app depends on image recognition, artificial intelligence, machine learning and deep learning algorithms to diagnose plant health [80, 83, 79, 81, 68]. The user (such as a farmer or plant expert) takes a photo of a plant in the field. Then, the user either directly uploads the image to the PEAT servers or posts it whenever an internet connection becomes available. Once uploaded, the app automatically records the time of capture and the location. The Plantix app quickly analyses the image through a deep neural network (Plantix-DNN) by using multiple ConvNets using images available on the server [80, 83, 79, 81]. One network in the ConvNets is to check the object in the image, and if there is no plant in it, the photo is deleted. One network defines crop type (name), and another one defines the health condition [80, 83, 79, 81]. The app compares a user's image with a vast database of high-resolution images of various crops and diseases and provides accurate diagnosis [68]. Then, the most similar crop disease to the user image according to a highest softmax probability (top-1 prediction) is displayed to the user. Other predicted lower probabilities, such as top-2 prediction, can be presented in lower positions, which can be used solely for internal purposes related to research and the company itself. Therefore, the app can be used as a decision support tool. The Plantix app provides users with the crop type, predicts plant diseases, and offers information on biological or chemical treatments to mitigate the disease. It also gives valuable information on preventive measures to avoid the plant disease occurrences or at least discover them at an early stage [80, 83, 79, 81].

Today, the Plantix app is available in 150 countries and in 18 different languages.

In 2016, 25,000 farmers used the application in Germany, and PEAT published the app to a global level in Brazil, Tunisia, Kenya and India. To date, it has been downloaded over 15 million times and millions of farmers use it. Day by day the number of images grows, as PEAT's database is expanded through people who upload pictures on the app to find the answer to their plant problems and cooperate with PEAT's private and public partners. By April 2021, Plantix app diagnosed over 35 million pictures. Plantix app has offices in Berlin, Hyderabad and Indore, and this free agricultural app becomes the most has been used around the World [83, 84, 85, 81].

In India, the Plantix app has been used since 2016, in cooperation with the Centre for Agriculture and Bioscience International (CABI), and the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT). PEAT has collaborated with ICRISAT as a partner that help to extend PEAT's database to include Indian crops and diseases that differ from other countries. The early plan of this cooperation was to assisst farmers in Telangana and Andhra Pradesh. In early 2017, ICRISAT organised several workshops in six districts of Andhra Pradesh and Telangana to teach farmers how to use Plantix app. The app is available in Indian regional languages, and the first Indian regional languages were Telugu and Hindi. In India 8.6 million samples were uploaded onto the app between 2017 and 2018. Most of these submissions were during the harvest of the Kharif season which between Septamber and November [79, 83, 86, 81, 82].

Although Plantix app is a popular and highly rated app for detecting and managing plant diseases, it also has some potential disadvantages. One disadvantage of the app is that it may not be completely accurate at identifying problems in plants. Hence, there is a risk of misidentification either false positives or false negativies. A study by Siddiqua et al. (2022) evaluated various apps (17 apps) for detecting and managing plant diseases based on a set of predefined functionalities. It was found that no single app encompassed all seven functionalities, which include disease detection, plant identification, disease severity estimation, treatment suggestions, community support, identification of affected plant parts, and plant coverage. While the study highlighted the Plantix app as one of the top options for disease detection and management, it did not conclude that it was the best app overall. Nevertheless, among the 17 apps reviewed, Plantix emerged as the most popular. The study noted that Plantix received high ratings for its automated plant identification and disease detection capabilities, as well as for the expert recommendations it provides [68].

Furthermore, another disadvantage is that the app is not capable of detecting all types of plant species [68]. In a study conducted in India by Wang, Di Tommaso and others, it was found that the overall accuracy of Plantix-DNN to define Kharif crop types of 72,494 samples was 97 percent, while for maize crop it was more than 80 percent [79, 83, 86, 81, 82]. Moreover, an additional drawback is the sampling bias, whereby farmers without smartphones and internet access cannot participate in data collection. Additionally, volunteers' efforts fluctuated over time because of seasonal patterns or declining commitments. Based on this challenge, control of data collection or data analysis should be considered [79, 87, 88].

1.4 Statistical and mathematical modelling

This thesis employs a comprehensive framework of mathematical and statistical analysis to analyse the Plantix app data about FAW in maize in India, and the expert data about the EAB in the ash plants in the USA. The details, including literature reviews and research gaps of these methods are presented in the relevant chapters, Chapter 3 to Chapter 5 cover the statistical analysis of the Plantix app data, and Chapter 6 deals with the EAB.

1.4.1 Statistical and mathematical modelling of the Plantix app data

The Plantix app dataset was provided by PEAT GmbH. This dataset consists of notifications in Indian maize crop, including those diagnosed as infested with FAW and those without FAW infestation. The timeframe of the dataset is from January 1, 2018 until December 31, 2019. Within this period, there were 138359 data points from 631 districts. The dataset includs the time stamp, longitude and latitude GPS coordinates and the results of the machine learning as categorical and numerical variables. The categorical variables were "top-1 pathogen name" and "top-2 pathogen name". These two variables define the detection status of each maize diagnosis as either healthy (indicating the maize is in good health), infested with FAW, infested with pathogens, affected by other pests, or suffering from nutrient deficiencies. The numerical variables were "top-1 similarity" and "top-2 similarity", representing positive integer values ranging from 2 to 100. A large similarity value indicated that the detection was more confident of reflecting the real status of the diagnosed maize. It is noteworthy to mention the

distinction between top-1 and top-2. The DNN generates several similarity scores with corresponding health conditions for each tested crop where the total sum of these similarities equals 100% (or a probability of 1). Further, under the top-1, the highest value is recorded, while top-2 includes the next highest value.

Although the Plantix app has been widely discussed in agricultural research due to its usefulness and practical advantages [68, 89, 90], a limited number of studies have used the Plantix app dataset for studying diagnostic accuracy or understanding plant pest and disease behavior over space and time. Akinyemi et al. [91] evaluated the accuracy of the Plantix app in diagnosing FAW damage in maize in Nigeria using only ten images from infested maize plants and ten images from healthy maize plants. The study results finds that the Plantix app reached 100% accuracy in recognising the FAW symptoms as well as the healthy maize. However, the very small sample size of only ten images per condition reduces the reliability of generalising these findings regarding the Plantix app accuracy in diagnosing healthy or FAW damage in maize.

Hampf et al. [80] used the Plantix app to detect pests and diseases in maize and soybean crops in the southern Amazon, Brazil. The diagnosed crops with the top-1 similarity rate greater than 0.5 (50%) were selected to be as true positive cases, where diagnosed crops below this threshold were excluded to reduce the likelihood of misclassification and enhance the overall quality of the dataset. Then, the selected high-confidence observations were used to assess the Plantix app accuracy by comparing the diagnostic app results with actual conditions (gold standard). The sensitivity of the Plantix app was reported as 91.51% for maize, although it was lower for soybean diseases and pests. However, using a cut-off value of 50% may not always represent a good hypothesis, where the selected observations may include cases with similarity rate greater than 50% that are incorrect or less reliable A more statistically rigorous approach, such as a statistical classification framework could provide a more robust criterion for filtering observations.

Integrating the Plantix app data with statistical and mathematical modelling leads to explainable, robust, transforming big data into structured knowledge and enabling better decision-making. Although the diagnostic accuracy of the app remains unknown [68], estimating its accuracy was a critical gap that needed to be filled. Low accuracy, combined with a large number of users, can increase the risk of providing incorrect diagnoses, which may lead to inappropriate treatment, reduce plant quality, and fail to

stop the spread of pests

Accordingly, to effectively estimate the app accuracy, the app's internal database needs to be classified into true and false diagnoses. Moreover, it is fundamental to study a single pest within a specific crop. This is because it is unrealistic to expect one test to have uniform accuracy for all pests and crops. Hence, the generalisation of the accuracy would not provide useful information for the user. Therefore, Fall Armyworm data in maize crop in India was selected data for further analysis.

This choice was not only due to its importance of FAW and maize but also for pragmatic reasons. Firstly, the Plantix app recorded thousands of spatio-temporal data points related to maize for diagnosing Fall Armyworm, providing a substantial dataset for statistical analysis. Furthermore, this work is part of a collaboration between ICRISAT, which focuses on Fall Armyworm in maize, and PEAT. PEAT aims to learn how their data can support research on Fall Armyworm in India. These factors collectively made the Fall Armyworm data in maize the most suitable and strategically significant choice for this study.

Direct assessment of the data can be done by comparing with a gold standard data. A gold standard classifies perfectly between true and false diagnostic results of the target feature for each unit and estimates unbiasedly the accuracy of diagnostic test [92]. However, in this study, there is no accessible or feasible gold standard dataset for the Plantix app data. Establishing a gold standard dataset for the Plantix app data poses significant challenges. Firstly, it would require accessing the original images stored on the Plantix app server, which is difficult due to the huge dataset size. Secondly, since maize is a seasonal crop and the data were collected during the 2018–2019 growing seasons, it is no longer feasible to revisit the fields for expert verification of the diagnoses. These limitations emphasise the practical limitations in defining a gold standard for the Plantix app dataset.

Therefore, statistical parametric and nonparametric methods need to be applied to classify true and false diagnostic outcomes of the Plantix app and to estimate the gold standard list [69]. When gold standard is absent, latent class models, mixture models, logistic regression and receiver-operating characteristic (ROC) curve are statistical classification methods that have widely been used [93, 69, 94], and are some of them are implemented here. In this study, internal evidence from the data itself was used to classify observations into true and false categories using mixture model. Further,

since the goal is not to use the data for prediction but rather to understand its behavior during the time frame of the data, there is no need to test the validation of the data. The focus is on analysing the data's patterns and trends.

After the classifiction and extracting the true and false data, they can be used to estimate the app accuracy. To the best of our knowledge, ours is the first research that estimates the sensitivity and specificity of the Plantix app, for detecting FAW in maize. In addition, we believe that in the existing literature, there has not been a study that statistically evaluates the Plantix app accuracy in any other insects. The estimate, in this thesis, is made within a single dataset and under the absence of a gold standard or alternative methods for comparison. Therefore, a statistical framework is introduced to handle these challenges by integrating meta-analysis with Bayesien statistics. Finally, the FAW prevalence over space and time was estimated, using four different statistical methods to find a more accurate estimation.

1.4.2 Statistical and mathematical modelling of the expert data

In addition to the AI diagnostic data, a traditional method of expert diagnostic data was also analysed in this thesis. The expert data about emerald ash borer (EAB) invasive insect in the ash species (*Fraxinus spp.*) from 2002 to 2020 in the USA counties was used. This data was provided by United States Department of Agriculture Animal and Plant Health Inspection Service (USDA APHISUS). However, a specific challenge with this data is that it only includes an initial true positive case from each observed infested county, neglecting subsequent occurrences. This limitation means that density or prevalence cannot be estimated due to the single value per county. However, the spatial-temporal distribution of the EAB can still be modeled to provide insights into its spread and behavior under the influence of climatic and non-climatic factors.

Numerous studies of EAB have employed a variety of mathematical and statistical models to understand the insect's spread, either at local or reigonal scale. Related models have meticulously accounted for one or more influential factors in the spread of EAB. These factors encompasse ash tree availability or density, EAB presence-absence [46, 58] or abundance [95], the distance between infested and uninfested areas, as well as human and climate-related influences [96]. Most of these studies are focussed on North America, and primarily spatial scales, ranging from cells to subcounties to counties [96, 43]. The majority of studies consider temporal scale, within year to understand

EAB either in North America or Europe [97, 96, 43]. The results of these studies have provided valuable insights into spread of EAB as well as next scientific research.

Moreover, ash tree intensity is critical since EAB is more likely to spread in a high intensity area of ash trees [41, 98, 58]. Further, most studies focus on presence-absence data compared with fewer studies of EAB abundance [95]. Climate represents another essential class of variable to be considered when understanding and forecasting the potential dynamics of EAB. The number of growing degree days (GDD) has a significant impact on survival of insects, and determining the possible dispersal landscape of EAB adults [97, 99]. For example, low heat availability can limit the suitable area for EAB establishment. According to Orlova-Bienkowskaja [97] EAB adults are unable to colonise an area where its growing degree day accumulation across the year it lower than 700 degree-days.

Additionally, a number of previous efforts have highlighted the importance of long-distance dispersal in predicting the extent of EAB invasion [43, 100, 58, 101], because anthropogenic factors have become a weighty contributor to the spread of EAB. Therefore, dispersal models have been widely applied in previous research as valuable tools for understanding the spread of EAB in a local and regional areas. To account for both natural spread of EAB and human-mediated long-distance dispersal, most studies utilise a dispersal kernel [41]. In EAB modeling, the negative exponential dispersal kernel (exponential decay function) was most commonly applied [41]. The negative exponential kernel was used to model EAB spread in North America by Muirhead et al. (2006) [43] and Kovacs et al. (2010) [58] and in Europe by Orlova-Bienkowskaja et al. (2018) [46]. In addition to negative exponential kernel function, Orlova-Bienkowskaja et al. (2018) [46] applied also a Cauchy (fat-tailed) model and normal kernel function to predict the EAB spread. They implemented these dispersal functions in European Russia and neighboring countries. They concluded that the Cauchy model is the most appropriate for understanding and predicting the EAB spread, when the pairwise distance between locations is greater than 200 km [46]. Note that, Bienkowskaja et al. [46] did not consider a power law kernel, where the best model, Cauchy, can be closely approximated by a power law kernel.

Although a range of variables and factors were addressed in previous research using models of varying complexity, all of them enhanced the ongoing scientific research, forestry knowledge and governmental information. For example, some models simply

considered a distance based kernel function as a factor to predict the EAB spread [46, 58, 43]. Some used only climatic variables in climate-based ecological niche models, to predict the EAB spread by determining the suitable climate for EAB in North America [102, 45]. On the other hand, Prasad et al. [101] used a complex mechanistic model to estimate and forecast the EAB spread in Ohio, USA. Their model is a spatially explicit cell-based and a combination of two components: i) a short spread model, which reflected the EAB natural flight, and ii) a long distance model, which simulated the long distance spread due to human practices. The factors considered important in the long distance model are traffic density on major roads, wood products weights, population density and campgrounds.

The generic colonisation–dispersal model proposed by Catterall et al. [103] is a stochastic spatio-temporal model. It is a generic model designed to be applied when investigating a diverse range of invasive species in a different geographic regions. It aims to estimate and predict the spread of a specie through a space and over time. Therefore, here the model is adapted to study EAB spread in the USA, considering climatic (anuall average of growing degree day) [97], non climatic (ash density habitat) conditions, and dispersal mechanisms. To our knowledge this study is the first to apply the generic colonisation–dispersal model in EAB study on the whole of the USA. Additionally, while the previous studies in the USA included temperature as a climate variable in estimating EAB spread, this research replaces temperature with GDD. Chapter 6 bridges these gaps, and concluded that the significance of incorporating climatic and non-climatic factors as well as dispersal mechanisms.

1.5 Thesis scope

As discussed earlier, one of the main steps in managing insect pests effectively needs to early pest detection, and reliable spread assessment, and this can be achieved through statistical analysis. The Plantix app data on Fall Armyworm in maize crop in India between 2018 to 2019 provides early detection observations and a large dataset. However, lack of gold standard and imperfect accuracy were also noted. In addition, the expert data for EAB in the USA from 2002 to 2020 includes early detection observations and assumes a perfect accuracy in diagnosing. Nevertheless, providing a first observation from each county limits the statistical analysis. Thus, this thesis aims to address these gaps. Statistical analysis has been used to understand FAW prevalence in India, after

the gold standard data is estimated and the app accuracy is determined. In addition, spatially extended model is developed for EAB spread, incorporating host location, environmental factors, and dispersal in order to define reliable spread factors. The thesis develops several decision-making tools for pest management. Towards achieving these aims, five research objectives have been established. In this section, these research objectives and thesis structure are briefly outlined

The first purpose of this research is to enhance the importance of the analysis of the AI data before using as a tool in important subjects. Therefore, four ojectives under this main goal were (i) to classify data into "true" and "false" observation, (ii) to measure the temporal spreading of the FAW, (iii) to evaluate the accuracy of the AI diagnostic tool, and (iv) to understand the spatio-temporal prevalence of the FAW.

The second objective is to use expert data about historical occurance records of EAB, geographic land cover, environmental and climatic factors, as well as natural and human-mediated dispersal of the EAB to understand a spatial-temporal patterns of EAB in the USA. Once the purposes mentioned above have been achieved, this research will have answers to the following main research questions.

1.5.1 Research objectives and questions

First objective: Conduct classification model to filter the Plantix app data.

- How should researchers process imperfect data before analysis?
- Are all notifications provided by the Plantix app for diagnosing a Fall Armyworm pest in maize crop accurate?
- If not, how can the true notifications be distinguished from the false ones?

Second objective: Visualise and analyse the temporal variations in pest intensity within the true notifications.

- How does the intensity of the pest vary over time in the true notifications identified by the classification model?
- Are there any seasonal patterns or trends in the pest intensity data over time?

Third objective: Evaluate the accuracy of the Plantix app in detecting specific pest in a particular crop.

- How sensitive is the app in detecting truly FAW observation?
- What is the specificity of the app in accurately tested maize observations that truly not infested with FAW?

Fourth objective: Estimate the spatial-temporal prevalence of FAW infestations in maize filed in India. Highlight impacted regions

- What is the seasonal spatial-temporal maps of FAW infestations in India during study period 2018-2019?
- How is the direction of FAW spatial distribution changed over season in India?

Fifth objective: Understand EAB biology dynamics and relations with host plants and climate.

- What is the current spatial-temporal map of EAB infestations in the USA county level?
- What role do land cover and climatic factor, play in influencing the abundance and distribution of the EAB?
- In what extent does human-mediated dispersal role contribute to EAB colonisation over time?
- What is the estimated model to imitate EAB spread over the study period in the USA?

The value of this thesis is in providing suggestions and recommendations for foresters, farmers, researchers, and government officials who are responsible for controlling the direct or indirect impacts of alien insects. In addition, this research enhances our ability to assess spatiotemporal dynamics of non-indigenous insect species, their relation with the environment, and how provide advice on people can better manage and limit them in the future.

1.5.2 Thesis structure

The rest of the thesis is structured as follows

Chapter 2 provides general description of selected methods and models that are used throughout different chapters. **Chapter 3** applies classification model, bivariate

mixture model, to identify the True and False classification of both positive (Plantix app notifications with maize infested with FAW) and negative (Plantix app notifications in the absence of FAW) observations. The parameters of the model are estimated using the expectation-maximization (EM) algorithm. Additionally, this chapter visualises the intensity of FAW over the invision period (2018-2019), where *odds* metric is used. The *odds* is the ratio of the truly presence of FAW (True positive) compared to the truly absence of FAW (True negative). This chapter addresses first and second objectives.

Chapter 4 implements the Bayesian-meta analysis to accomplish third research objective. The chapter briefly discusses meta-analysis. The meta-analysis is a statistical technique usually utilised to estimate a pooled interested measure(s), integrating data from multiple independent studies. These studies perform similar methodology, while the studies can relate to diverse populations [104]. Chapter 4 incorporates the meta-analysis, assuming each state in India as an independent study, where each state having its own government and administrative structure. The model is bivariate generalised linear random effects model. The parameters of the model are estimated using the Markov Chain Monte Carlo sampling (MCMC) to take into account for the uncertainty. Further, the pooled sensitivity and specificity estimate in this chapter is used as input values in the Method-2 to estimate FAW prevalence in Chapter 5

Chapter 5 includes four different methods to select most appropriate modeling techniques for estimating FAW prevalence in Indian states during each maize seasons Kharif and Rabi, 2018 and 2019 (fourth objective). These methods range from simple to comprehensive. Method-1 is based on a frequentist technique, where FAW prevalence is calculated using a 2x2 diagnostic Table. Method-2 uses the Rogan–Gladen adjusted estimator with sensitivity and specificity values from Chapter 4. Method-3 uses Chapter 3 methodology to estimate sensitivity and specificity for each maize seasons between between 2018 to 2019. Then, the Rogan–Gladen adjusted estimator with sesonal sensitivity and specificity values is utlised. Method-4 relies on a comprehensive Bayesian meta-analysis with stochastic sensitivity and specificity. This technique is similar to previous chapter methodology but more informative. This is because it has additional likelihood functions and prior distributions, including priors for each parameter within the model (hierarchical prior).

In **Chapter 6**, the expert data about emerald ash borer (EAB) in the USA counties is used. A generic colonisation-dispersal model proposed by Stephen Catterall and his

colleagues (2012) [103] was adapted to develop a continental level models for the spread of EAB in the USA, fitting this to the best available data, quantifying the uncertainty in these models and their predictions and then assessing their performance in tracking spread of EAB over two decades. These models consider climatic (annual average of growing degree day) and non climatic (ash density habitat) conditions, and dispersal mechanisms. They differ in using different functions and scaling methods for the climatic factors and two different dispersal kernel functions to understand the natural of the EAB dispersal sistance. **Chapter 7** concludes the thesis by discussing the contributions to knowledge and potential future directions that can further develop upon this work.

Chapter 2

Overview of established methods and models for diagnostic test classification

The present chapter outlines a comprehensive overview of the general methods that were derived from the literature. These methods were utilised in the subsequent chapters, with related detailed applications were explained.

2.1 The 2x2 diagnostic test table

A 2x2 diagnostic test table is a contingency table that outlines the reliability and accuracy of a diagnostic test [105], where the outcomes can be represented as binary (dichotomous) outputs which are either a positive (P) or negative (N). The positive outcome means that an studied observation has the target characteristic, while a negative finding indicates the absence of the target characteristic in an observed event [106]. The P and N outcome results are grouped into four cells within the 2x2 diagnostic test table. Each cell represents a frequency count as well as a unique characteristic that describes the accuracy of a test or classification algorithm. One of the best method to define the frequency of the cell in the 2x2 diagnostic test is a gold standard, which is a list of the actual target characteristic of the studied observation [105]. Therefore, the four outcomes of the 2x2 diagnostic test table are:

- True positive (TP): This cell reflects the number of instances where the diagnostic method correctly predicts the target condition (positive).

- False positive (FP): The number of observations where the diagnostic tool incorrectly signs a target characteristic as present (positive), whlie the observation does not have the target characteristic.
- True negative (TN): The number of observations in which a diagnostic tool correctly identifies the absence of a targeted characteristic (negative) when the observation actually does not have the target characteristic.
- False negative (FN): The number of cases wherein the diagnostic method result is negative, while the observation actually has the condition being tested for [106, 105].

Table 2.1: A 2x2 diagnostic test table, counts of binary classifier model vs. target characteristic.

Binary classifier model	Target characteristic	
	Present	Absent
Positive	TP	FP
Negative	FN	TN

The 2x2 diagnostic test table is a simple and effective technique that can be used to calculate sensitivity and specificity in order to assess the perofomrance of the dichotomous diagnostic tool. Sensitivity (true positive rate, TPR) is the probability of the true positive, which reflects the ability of the diagnostic tool to identify the observations that truly have the target characteristic. On the other hand, specificity (true negative rate, TNR) is defined as the probability of the true negative, which indicates the ability of the diagnostic tool to identify the observations that truly do not own the target characteristic [106, 105].

$$\text{sensitivity} = Se = \frac{\text{TP}}{\text{TP} + \text{FN}} \qquad \text{specificity} = Sp = \frac{\text{TN}}{\text{TN} + \text{FP}}$$

Sensitivity (Se) has a value that ranges between 0 and 1. A sensitivity value of $Se = 0$ indicates that there are no true positives, which signifies a complete failure of the test to identify any actual positive cases. In this instance, all positive events are incorrectly classified as negative (FN), leading to a total lack of detection. Conversely, a sensitivity value of $Se = 1$ means that there are no false negatives, indicating that the test successfully identifies all actual positive cases. This represents an ideal situation

where every positive instance is correctly recognised, and there are no missed detections. Another significant value is $Se = 0.5$, which corresponds to a situation where the test has a balanced performance in identifying true positives and false negatives. In this case, the test correctly identifies half of the actual positive cases while failing to detect the other half.

Similarly, the specificity (Sp) also ranges from 0 to 1. A specificity value of Sp=0 indicates that the test fails to identify any true negatives (TN), meaning that all negative instances are incorrectly classified as positive (FP). This results in a complete failure to recognise the absence of the condition being tested for. On the other hand, a specificity value of $Sp = 1$ signifies that there are no false positives (FP), meaning that the test accurately identifies all negative cases (TN) without mistakenly labeling them as positive.

However, the gold standard may not be up to date, adequate or cost effective for diagnostic test. As a result, the constructing of the 2x2 table can be challenging. In the next chapter we will introduce a new method that allows the classification without the gold standard. This method uses a number of strategies, one of which is mixture of two distributions [94].

2.2 Mixture of two normal distributions with EM algorithm

Let X be a random variable, reflecting the outcome of a real-life application. X may not be unimodal, where it can be multimodal. This suggests that X is a mix of several distributions, where each distribution has a single mode and corresponds to a simple parametric distribution. The distributions may belong to the same family with different parameter values, or different familes [107]. The procedure of modelling several components in a single model is called a mixture model, and it is built of two essential variables: observed variable and unobserved (latent) variable. The observed variable is a set of observable and collected units, while the latent variable is hidden where the data does not directly consider. It usually represents grouping within the data. The latent variable's occurrence can be discovered by having more than one mode in the observed variable; in classification language each mode represents a group. To sum up, the mixture model is composed of a sum of distributions with mixing weights (propor-

tions), where mixing proportions are non-negative and the sum of them equals one. The mixture model can be formed from continuous or discrete distributions and from at least two distributions [108, 107]. The subsequent lines provide an explanation of the mixture model based on existing literature [109, 110, 111, 108].

To formulate the mixture model, the latent variable $\boldsymbol{z}$ takes discrete values, $\{1, 2, ..., K\}$, with K the number of observed parametric forms. The observed variable is $\boldsymbol{x}$ and each data point x is linked to a single value of $\boldsymbol{z}$,which determines the component from which it is generated. However, when analysing the data, x may have a likelihood of coming from multiple components, with different probabilities for each. Therefore, the key expression of the mixture model is built by the following process:

- Firstly we sample $\boldsymbol{z}$ from its distribution $p(z)$. This step determines which component of the mixture the observation will come from. $p(z)$ is a prior distribution that can be derived from a multinomial distribution, $z \sim Multinomial(\boldsymbol{\omega})$, where $\boldsymbol{\omega}$ is a vector of mixture proportions of size (also called mixture weights or prior probabilities) K. For example, K-means clustering can be used to initialize these proportions ω.

- Then, we sample $\boldsymbol{x}$ given z (i.e. from the conditional distribution $p(x \mid z)$). This step produces the actual data point based on the selected component. $p(x \mid z)$ is a parametric distribution. It is a conditional probability and called a mixture component. It can be a continuous or discrete distribution.

- Hence, the joint probability of the mixture model is given by:

$$p(\boldsymbol{z}, \boldsymbol{x}) = p(z)\, p(x \mid z).$$

- Finally, summing of the joint probability over all possible values K of the latent variable z is the mixture model. This can be achieved using the law of total probability (i.e. summing the probability of each observation x over all possible z, weighted by each z probability $p(z)$). This is also knwon as the marginal probability distribution $f(x)$ of the x is

$$\begin{aligned} f(x) &= \sum_{k=1}^{K} p(\boldsymbol{z}, \boldsymbol{x}) \\ &= \sum_{k=1}^{K} p(\boldsymbol{z} = k)\, p(\boldsymbol{x} \mid \boldsymbol{z} = k) \\ &= \sum_{k=1}^{K} \omega_k\, f_k\left(\boldsymbol{x}; \boldsymbol{\theta}_k\right) \end{aligned}$$

The $f(x)$ here is the weighted average of the mixture components. It gives the overall probability of x regardless of the value of z. The index K is the number of finite mixture components equivalent to the number of modes, $2 \leq K$. Here, $\omega_k \in \boldsymbol{\omega}$ satisfies the conditions $0 \leqslant \omega_k \leqslant 1$ and $\sum_{k=1}^{K} \omega_k = 1$. $\boldsymbol{x} = [x_1, x_2, ..., x_n]^T$is a vector of size n, and we assumes that each point is independent and identically distributed, $x_i \sim f_k(x_i; \theta_k)$. $f_k\left(\boldsymbol{x}; \boldsymbol{\theta}_k\right)$ are univariate distributions with a set of parameter $\boldsymbol{\theta}_k$. To simplify, let $\boldsymbol{\theta} = (\omega_1, ..., \omega_K, \theta_1, ..., \theta_K)$ a set of the all parameters in the $f\left(\boldsymbol{x}\right)$. Then, the likelihood of $f\left(\boldsymbol{x}\right)$ is:

$$L\left(\boldsymbol{\theta} \mid \boldsymbol{x}\right) = \prod_{i=1}^{n} \sum_{k=1}^{K} p(z_i = k)\, p(z_i, x_i) = \prod_{i=1}^{n} \sum_{k=1}^{K} \omega_k f_k\left(x_i; \theta_k\right).$$

and the log likelihood is:

$$\ell\left(\boldsymbol{\theta}\right) = \sum_{i=1}^{n} \log \left[\sum_{k=1}^{K} \omega_k f_k\left(x_i; \theta_k\right) \right].$$

To identify the points where the $\ell\left(\boldsymbol{\theta}\right)$ reaches its maximum using maximum likelihood estimation (MLE), sets the derivative to be zero. However, the summation inside the logarithm cannot be solved analytically for each parameter, since it gives a complicated expression for MLE. For example: to take the derivative with respect to θ_k,

$$\begin{aligned}
\frac{d\ell\left(\boldsymbol{\theta}\right)}{d\theta_k} &= \sum_{i=1}^{n} \frac{d}{d\theta_k} \log\left[\sum_{k=1}^{K} \omega_k f_k\left(x_i;\theta_k\right)\right] \\
&= \sum_{i=1}^{n} \frac{1}{\sum_{k=1}^{K} \omega_k f_k\left(x_i;\theta_k\right)} \frac{d}{d\theta_k}\left[\sum_{k=1}^{K} \omega_k f_k\left(x_i;\theta_k\right)\right] \text{ using chain rule} \\
&= \sum_{i=1}^{n} \frac{\sum_{k=1}^{K} \frac{d}{d\theta_k}\left[\omega_k f_k\left(x_i;\theta_k\right)\right]}{\sum_{k=1}^{K} \omega_k f_k\left(x_i;\theta_k\right)} \\
&= \sum_{i=1}^{n} \sum_{k=1}^{K} \frac{\frac{d}{d\theta_k}\left[\omega_k f_k\left(x_i;\theta_k\right)\right]}{\sum_{k=1}^{K} \omega_k f\left(x_i;\theta_k\right)} = 0.
\end{aligned}$$

It is not possible to derive an analytical solution for θ_k. This is because these parameters are not independent of each other, while we should estimate them simultaneously. However, if the latent variable $\boldsymbol{z}$ is known, it simply gathers all x_i in each $\boldsymbol{z} = k$ and simply uses the MLE. To help us to find the MLE when $\boldsymbol{z}$ is unknown, the expectation-maximisation (EM) algorithm can be used.

The expectation-maximisation (EM) algorithm is a method to estimate probability distributions parameters with the latent variables present by performing the MLE. The EM algorithm was published by Demster, Laird, and Rubin in 1977. It can be applied iteratively as an approach that cycles between two steps. The first one is an expectation-step or E-step, because it computes the expected value of the latent variable ($\boldsymbol{z}$) given the data (x) and the current values of the parameters ($\boldsymbol{\theta}^{(t)}$). In other sense, it calculates the posterior distribution $p(z|x, \boldsymbol{\theta}^{(t)})$ of the latent variable ($\boldsymbol{z}$).

The second step is a maximisation-step or M-step, which attempts to update the parameter values given the current posterior probabilities (E-step). As a result, this step maximises the likelihood of the observed data given the estimated posterior probabilities. Hence, M-step generates a new value for the parameters, which are used for the next iteration of the algorithm. To formulate the EM algorithm, we assume that we have the complete data log-likelihood, including both the observed data and the latent variable. Then, we reformulate the derivate of the above marginal log likelihood as following:

$$\begin{aligned}
\frac{d\ell\left(\boldsymbol{\theta}\right)}{d\theta_k} &= \sum_{i=1}^{n} \frac{d}{d\theta_k} \log\left[\sum_{k=1}^{K} \omega_k f_k\left(x_i;\theta_k\right)\right] \\
&= \sum_{i=1}^{n} \frac{1}{\sum_{k=1}^{K} \omega_k f_k\left(x_i;\theta_k\right)} \frac{d}{d\theta_k}\left[\sum_{k=1}^{K} \omega_k f_k\left(x_i;\theta_k\right)\right] \text{ using chain rule} \qquad (2.1)
\end{aligned}$$

The term $\frac{d}{d\theta_k}\left[\sum_{k=1}^{K}\omega_k f_k\left(x_i;\theta_k\right)\right]=\sum_{k=1}^{K}\omega_k f_k\left(x_i;\theta_k\right)\frac{d}{d\theta_k}\log\left[\sum_{k=1}^{K}\omega_k f_k\left(x_i;\theta_k\right)\right]$.
Now rewrite the equation 2.1 again

$$\begin{aligned}
\frac{d\ell\left(\boldsymbol{\theta}\right)}{d\theta_k} =& \sum_{i=1}^{n}\frac{\sum_{k=1}^{K}\omega_k f_k\left(x_i;\theta_k\right)}{\sum_{k=1}^{K}\omega_k f_k\left(x_i;\theta_k\right)}\frac{d}{d\theta_k}\log\left[\sum_{k=1}^{K}\omega_k f_k\left(x_i;\theta_k\right)\right]\\
=& \sum_{i=1}^{n}\sum_{k=1}^{K}\left[\frac{\omega_k f_k\left(x_i;\theta_k\right)}{\sum_{k=1}^{K}\omega_k f_k\left(x_i;\theta_k\right)}\right]\frac{d}{d\theta_k}\log\left[\omega_k f_k\left(x_i;\theta_k\right)\right]\\
=& \sum_{i=1}^{n}\sum_{k=1}^{K}\widehat{\gamma}_{ik}\left[\frac{d}{d\theta_k}\log\omega_k f_k\left(x_i;\theta_k\right)\right]\\
=& \sum_{i=1}^{n}E_{p\left(z_i=k|x_i,\theta^{(t)}\right)}\left[\frac{d}{d\theta_k}\log\omega_k f_k\left(x_i;\theta_k\right)\right]\\
=& Q(\boldsymbol{\theta}\mid\boldsymbol{\theta}^{(t)}).
\end{aligned}$$

$\widehat{\gamma}_{ik}$ is called responsibility and it indicates how strongly each data point belongs to each component. Therefore, Hence, $Q(\boldsymbol{\theta}\mid\boldsymbol{\theta}^{(t)})$ is the expected complete data log-likelihood. Consequently, the EM algorithm procedure is as follows:

- Initially, determines the number of the mixture components K,
- Then, defines the initial values of all possible parameters.
- After that, iteratively refines the mixture components based on two steps
  - Expectation step: compuates the $\widehat{\gamma}_{ik}$
  - Maximisation step: compuates the ML parameters given these $\widehat{\gamma}_{ik}$

$$\arg\max_{\theta}\sum_{i=1}^{n}\sum_{k=1}^{K}\widehat{\gamma}_{ik}\left[\log\omega_k f_k\left(x_i;\theta_k\right)\right]$$

- the E-step and the M-step are repeated alternately until the difference $l\left(\boldsymbol{\theta}^{(t+1)}\right)-l\left(\boldsymbol{\theta}^{(t)}\right)<\delta$,where δ is small value. In this thesis, we consideres $\delta=1\times10^{-6}$.

In this thesis, we will assume $K=2$. We will also only consider $f_k\left(x_i;\theta_k\right)$to be normal. Hencc, the mixture of two-one dimensional normal distributions according to[109, 110, 111, 108]) can be written as:

$$f(x_1, x_2..., x_n; \mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \omega) = \prod_{i=1}^{n} \left[(1-\omega) \frac{1}{\sqrt{2\pi\sigma_1^2}} \exp\left(-\sum_{i=1}^{n} \frac{(x_i - \mu_1)^2}{2\sigma_1^2} \right) + \omega \frac{1}{\sqrt{2\pi\sigma_2^2}} \exp\left(-\sum_{i=1}^{n} \frac{(x_i - \mu_2)^2}{2\sigma_2^2} \right) \right].$$

with the proportion of each component $\omega_1 = 1 - \omega$ and $\omega_2 = \omega$. The parameters of the each normal distribution is $\theta_1 = \{\mu_1, \sigma_1\}$and $\theta_2 = \{\mu_2, \sigma_2\}$. The responsibility function (E step) becomes

$$\widehat{\gamma}_{i2} = \frac{\widehat{\omega} f_2\left(x_i; \mu_2, \sigma_2^2\right)}{(1-\widehat{\omega}) f_1\left(x_i; \mu_1, \sigma_1^2\right) + \widehat{\omega} f_2\left(x_i; \mu_2, \sigma_2^2\right)} = \widehat{\gamma}_i. \tag{2.2}$$

and then,

$$\begin{aligned} Q\left(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \omega\right) &= \sum_{i=1}^{n} \left[\widehat{\gamma}_{i1} \log\left[(1-\widehat{\omega}) f_1\left(x_i; \mu_1, \sigma_1^2\right) \right] + \widehat{\gamma}_{i2} \log\left[\widehat{\omega} f_2\left(x_i; \mu_2, \sigma_2^2\right) \right] \right] \\ &= \sum_{i=1}^{n} \left((1-\widehat{\gamma}_i) \log(1-\omega) + (1-\widehat{\gamma}_i) \left[-\frac{1}{2}\log(2\pi) - \log\sigma_1 - \frac{(x_i-\mu_1)^2}{2\sigma_1^2} \right] \right. \\ &\quad \left. + \widehat{\gamma}_i \log\omega + \widehat{\gamma}_i \left[-\frac{1}{2}\log(2\pi) - \log\sigma_2 - \frac{(x_i-\mu_2)^2}{2\sigma_2^2} \right] \right). \end{aligned}$$

Then, Q is differentiated with respect to each parameter

$$\begin{aligned} \frac{\partial Q}{\partial \mu_1} &= \sum_{i=1}^{n} \left((1-\widehat{\gamma}_i) \left(\frac{x_i - \mu_1}{\sigma_1^2} \right) \right) = 0 \\ \therefore\ \mu_1 &= \frac{\sum_{i=1}^{n} (1-\widehat{\gamma}_i) x_i}{\sum_{i=1}^{n} (1-\widehat{\gamma}_i)}. \qquad (2.3) \\ \frac{\partial Q}{\partial \theta_2} &= \sum_{i=1}^{n} \left(\widehat{\gamma}_i \left(\frac{x_i - \mu_2}{\sigma_2^2} \right) \right) = 0 \\ \therefore\ \mu_2 &= \frac{\sum_{i=1}^{n} \widehat{\gamma}_i x_i}{\sum_{i=1}^{n} \widehat{\gamma}_i}. \qquad (2.4) \end{aligned}$$

$$\begin{aligned}
\frac{\partial Q}{\partial \sigma_1} =& \sum_{i=1}^{n}\left[(1-\widehat{\gamma}_i)\left(\frac{-1}{\sigma_1}+\frac{(x_i-\mu_1)^2}{\sigma_1^3}\right)\right]=0 \\
\therefore\ \widehat{\sigma}_1^2 =& \frac{\sum_{i=1}^{n}(1-\widehat{\gamma}_i)(x_i-\widehat{\mu}_1)^2}{\sum_{i=1}^{n}(1-\widehat{\gamma}_i)}. \qquad (2.5) \\
\frac{\partial Q}{\partial \sigma_2} =& \sum_{i=1}^{n}\left[\widehat{\gamma}_i\left(\frac{-1}{\sigma_2}+\frac{(x_i-\mu_2)^2}{\sigma_2^3}\right)\right]=0 \\
\therefore\ \widehat{\sigma}_2^2 =& \frac{\sum_{i=1}^{n}\widehat{\gamma}_i(x_i-\widehat{\mu}_2)^2}{\sum_{i=1}^{n}(1-\widehat{\gamma}_i)}. \qquad (2.6)
\end{aligned}$$

to maximise the mixing probability ω,where it is the average of the responsibilities:

$$\widehat{\omega}_2 = \frac{1}{n}\sum_{i=1}^{n}\widehat{\gamma}_i$$

This method will be used in Chapter 3, Chapter 4 and Chapter 5.

2.3 A receiver operating characteristic (ROC) curve

The outcomes of classification models or diagnostic tools can be either binary (positive or negative), or on a continuous or ordinal scale. Thresholds can be used to categorize continuous or ordinal findings into binary outcomes when needed. For example, any findings above a certain level are assumed positives, while results below that level are assumed negatives [5]. By comparing these classifications with the true status, the True Positive Rate (TPR) and False Positive Rate (FPR) for each threshold can be calculated. Then, to obtain the performance of these models or tools, a receiver operating characteristic (ROC) can be used. ROC is a graphical curve that assesses the discriminative ability of a classification model. It plots TPR in y-axis against FPR in x-axis accross various thresholds, providing a comprehensive view of a test's performance [112, 5, 113]. Hence, the ROC curve provides more view of the model performance than the 2x2 diagnostic test table, which only assesses the performance of the model at a single threshold (cut-off) [5].

A diagonal straight line (chance line) draws between (0,0) and (1,1). The ROC curve position compared to the line reflects the model's performance. If the ROC curve is close to the diagonal line that indicates the classifier model randomly classifies the

data. In other words, the model cannot effectively distinguish between the positive and negative classes [112, 5]. If the ROC curve is under the diagonal straight line, the model performs worse than random guessing. It indicates that the model is unable to predict the negatives. The model performance increases as it moves to the top left corner (0,1) [5].

The ROC curve is very popular in many scientific areas such as radiology and epidemiology [113]. It is used to (i) assess the ability of the classification model or tool to distinguish between two classes truly; (ii) to determine an optimal threshold that achieves the maximum difference between the two target classes; (iii) and to compare between two or more classification models or tools for the same application [112].

The ROC curve can be used to measure the model accuracy and summarized in a single value by finding the area under curve (AUC). It is a dimensionless value that indicates the performance of the classifier model [114]. The AUC value ranges between 0 and 1. As the AUC value moves closer to 1, it indicates a better discriminatory ability of the model. [5]. Moreover, AUC can be used to compare the performance of different classification models in the same data, where the highest AUC value means the best model [112]. However, while AUC is widely used, it is not always the most suitable measure, especially in scenarios with highly imbalanced data (low prevalence). In such cases, a high AUC value may correspond to a model with low sensitivity, meaning it performs well at identifying negatives but poorly at detecting positives [115]. Thus, while the Table below summarises the general interpretation of AUC values, it is important to consider the context in which the model is applied:

Table 2.2: AUC Value and Model Performance [5].

AUC value	Interpretation
0.5	random
0.7-0.8	acceptable
0.8-0.9	excellent
0.9-1	outstanding

In addition, the ROC value can guide the selection of an optimal threshold for classification. The optimal threshold allows the construction of the 2×2 diagnostic table and to assess the model performence. The square of the distance and Youden index are two mathematical approaches to calculate the optimal threshold. The square of the distance calculates the squared distance between (0,1) and any order pairs on the

ROC curve. The (0,1) point is upper left hand corner of the ROC curve. It represents a perfect classification, because the TPR=1 and the FPR=0. In other words, it indicates that the model perfectly classifies all positive and negative observations. Therefore, the square of the distance can be written as following

$$\mathrm{d}^2 = (1 - \mathrm{TPR})^2 + \mathrm{FPR}^2.$$

The ROC point that has the minimum value of d^2 is considered as the optimal cut-off point and all predicted observations not less than TPR are identified as positives and the rest as negatives [112].

The Youden index searches the maximum vertical distance between chance line and a point in the ROC curve [112, 113]. Therefore, the index can be defined as the maximum difference between TPR and FPR and can be written as

$$\mathrm{Youdon\,Index} = J = \max\left(\mathrm{TPR} - \mathrm{FPR}\right).$$

The index ranges between 0 and 1, where $J = 1$ means the maximum difference between positive and negative groups. Moreover, it means the model can perfectly separate the two classes. Therefore, as J value is closer to 1, the ability of model to classify increases [113]. These metrics will be used in Chapter 6.

2.4 Growing degree day (GDD)

Growing degree day (GDD) is used in Chapter 6. It will serve as a climatic variable to investigate its association with the colonisation behavior of EAB. GDD is a metric used in agriculture and entomology to quantify the growth and development of plants and insects [116, 117]. It indicates the amount of accumulated heat above the threshold temperature during a specific time frame. The GDD is calculated by summing the values of the difference between the average daily temperature and a base temperature (threshold temperature).

The base temperature is the minimum temperature for an organism to grow or develop, below this value the plant growth or other biological processes cannot occur [117, 116]. Base temperatures differ among organisms, and different organisms require different GDD to activate metabolic processes [117, 116]. The base temperature can be experimentally and statistically determined, rather than physiologically, through

observing organism response to different temperatures [118].

The GDD formula can be written as follows:

$$\mathrm{GDD}\,(t_n) = \sum_{\mathrm{t}=t_1}^{t_n} (\ \mathrm{T_t} - \mathrm{T_b})\,.$$

where T_t is the average daily temperature, T_b is the base temperature and t_n is the total number of days.

GDD explores the time of the phenological dynamics, and determines the suitable area for growing plants or insect outbreaks. The GDD is a practical tool for farmers and agricultural scientists to track crop growth develpement such as crop flowering and predict harvesting time. In addition, entomologists have used the GDD to follow the insect pest development and predict outbreaks [118, 116]. The GDD gives an indication of the likelihood of insect activity where each insect requires a consistent GDD to reach definite life stages, such as egg hatch or adult flight [117, 118]. The GDD is a better method for modeling insects compared to using simple temperature measurements due to its biologically meaningful approach [116].

Additionally, GDD is an effective tool for understanding pest outbreaks over time and space and consequently for enhancing pest control [97, 99]. GDD, combined with the knowledge of climate change, forms a powerful tool to predict pest distributions. As a result, GDD can help pest managers and governments to limit and mitigate the outbreak [118]. It indicates which pests are likely to attack specific regions, determines the spatial distribution of damage, and the most likely time of invasion or infestation [119, 120]. In addition, GDD helps to determine the required type of control that is needed and the time of applying treatement. Therefore, it can help to reduce the risk of crop damage and economic losses [118].

2.5 Bayesian approach

In this thesis, Bayesian approach will be employed in Chapters 4, 5 and 6 to estimate model parameters. Chapter 4 and 5 will utilise Bayesian method within a meta-analysis framework for estimating diagnostic parameters, including sensitivity (Se), specificity (Sp), and prevalence. In Chapter 6, Bayesian method will be applied to estimate the parameters of models that help understanding the distribution of EAB under the influence of climatic and non-climatic factors, as well as the dispersal kernel of EAB.

In Bayesian statistics, any unknown parameter is associated with uncertainty, which can be represented by a probability distribution. A Bayesian approach allows the incorporation of prior knowledge and offers a more natural interpretation of uncertainty than a frequentist approach [121]. In Bayesian inference, observation data D provides information about a (set of) parameter(s), called likelihood, which is used to update a prior belief $f(\boldsymbol{\theta})$ about a (set of) parameter(s) into a posterior belief $f(\boldsymbol{\theta} \mid D)$ about a (set of) parameter(s) [122]. Formally, it can be explained by Bayes' theorem,

$$f(\boldsymbol{\theta} \mid D) = \frac{f(D \mid \boldsymbol{\theta}) f(\boldsymbol{\theta})}{f(D)} \propto f(D \mid \boldsymbol{\theta}) f(\boldsymbol{\theta}).$$

$f(D \mid \boldsymbol{\theta})$ is the likelihood for an observation of data D. $f(\boldsymbol{\theta})$ is prior probability distribution(s) for parameter(s) of interest. It represents the initial beliefs regarding each model parameters before any data is observed. The prior can be either informative or uninformative. Informative priors are shaping the distribution based on existing knowledge, such as previous analysis or literature. Without such information, the prior is usually selected in some logical way, such as uniform distribution over a realistic parameter range [121, 123, 103]. It can also be selected to have a minimal effect on the posterior distribution, leading to the likelihood to play a more role in explaining the posterior distribution.

Markov Chain Monte–Carlo (MCMC) has been an increasingly popular technique for estimating posterior distributions in Bayesian inference, which are often challenging to obtain analytically [121, 122]. It can handle complex models and high-dimensional parameter spaces, making it a versatile tool for a wide range of applications. Moreover, MCMC algorithms naturally incorporate prior information about the parameters of interest, which can improve the accuracy of inferences [124].

MCMC is a combination of two statistical approaches, a Markov Chain method and Monte–Carlo approach [125]. The former generates random samples until it converges to build up the target and optimal posterior distribution [124]. Every random sample serves as a stepping stone to produce the next random sample (hence the chain). The "Markov" attribute of the chain reflects the fact that while each new sample depends on the current one, they do not depend on any previous ones. After running the Markov Chain for a sufficient length of time, the samples' distributions converge to a stationary distribution reproducing the posterior distribution (if well-implemented) [122].

Monte Carlo is a way of generating random samples from the posterior distribu-

tion. Random samples are derived using a number of Monte Carlo methods, including rejection, importance, and Metropolis-Hastings sampling [122]. As the Markov Chain becomes larger, MCMC methods are capable of producing a sample of the Markov Chain size that represents the posterior distribution, from which to draw conclusions about the parameters. The sample can be used to estimate posterior probabilities, to calculate expected values and to obtain credible intervals for parameters [125, 122].

The MCMC method generates random samples from the posterior distribution of the parameters, after which inferences about the parameters are made. Statistics such as mean, median, and credible interval (CI) of a parameter can be estimated from the samples, reflecting the level of belief about its value [121].

2.5.1 MCMC conergence tools

A major step in the Bayesian analysis is to assess the convergence of the chains to the target posterior distributions. Four tools are frequently implented to evaluate the convergence of the MCMC simulations.

Trace (Time series) plots

The trace plot is the most frequently graphical method used as a convergence diagnostic tool in MCMC technique. It is a time series plot, and represents the values of the Markov chain over time in the state space. It shows the parameter values of the successive iterations on the y-axis against the corresponding iteration numbers on the x-axis. To evaluate the trace plot, the subsequent aspects should be considered:

- A converged chain reached stationary distibution, if minimal fluctuation (variation) occurs around a particular value in absence of visible trends or patterns. Hence, it represents appropriate mixing, generating the desired posterior distribution.
- The flat bits or lines in the trace plot indicate that the MCMC chain is not converging efficiently, and it is stuck in some part of the state space.
- The trace plot starts with initial samples, a "burn-in" period. In this period, the plot mostly shows poor convergence. Hence, the burn-in period should be discarded.

- Discarding burn-in samples may not be necessary, if the trace plot exhibits a consistent pattern throughout [126, 127].
- Multiple chains with various initial values are converge to the same distribution, if they are overlapping and appear indistinguishable [127].

Autocorrelation function (ACF) plot

The autocorrelation function (ACF) plot is another graphical statistical tool that is used to test and diagnoses the convergence of the MCMC simulations. The ACF can calculate the correlations between current estimated parameter and the previous estimated value based on the iteration number at lag k. In other words, the autocorrelation value at lag k, means the correlation between the current estimated value and the one k steps ealier. For example, the ACF value at lag k=4 measures the correlation between each current estimated parameter value and the value that was estimated at the 4 time steps earlier. However, at lag k=0 exhibits the relation with itself, as there are no previous samples, hence, the value is always one. To assess the ACF plot, the following should be taken into account:

- As the lag value k increases, the ACF plot should display a decay in the correlation. This is an indication that the MCMC simulations are well-mixed, the MCMC parameters samples are approximately independent and have converged to the target distribution.
  - Fast-mixing chain shows that the lag k autocorrelation values drop down to zero rapidly
  - Slow mixing chain means existence of correlation, and need to include a large number of samples to reach convergence [127, 126].

Gelman and Rubin diagnostic

Gelman and Rubin (1992) criteria is a single and dimensionless quantitative measure diagnostic tool, which is alternatively called the potential scale reduction factor (PSRF). The PSRF relies on analysis more than one chain that are simulated in parallel where each has a different initial samples. It compares the difference between the variance within the chains and the variance between the chains,

$$W = \frac{\sum_{i=1}^{m}\sum_{j=0}^{n-1}\left(X_{ij} - \bar{X}_i\right)^2}{m(n-1)} \quad \text{mean within chain variance}$$

$$\frac{B}{n} = \frac{\sum_{i=1}^{m}\left(\bar{X}_i - \bar{X}\right)^2}{(m-1)} \quad \text{variance between chains}$$

$$\hat{V} = \frac{(n-1)W + B}{n} \quad \text{pooled variance}$$

$$\hat{R} = \frac{\hat{V}}{W} = \frac{(n-1)W + B}{n} \times \frac{1}{W} = \left(\frac{n-1}{n}\right) + \frac{B}{nW}$$

where $i = 1, ..., m$ represents the index of each chain where the total number of chains is m. The index j where $j = 0, ..., n-1$ represents all samples within each chain. $\bar{X}_i$and $\bar{X}$is the average estimation for each chain and for all chains, respectively. The n in the term $\frac{B}{n}$ is to ensure consistency with the W value. The $\hat{V}$ is weighted average of the within and between variance chains. Finally, $\hat{R}$ is the point estimate of PSRF.

If $\hat{R}$ value approaches or equals one, it means that $W \approx B$. It indicates that chains are mixed and converge to the same distribution and with enough number of iterations [127, 126]. If $\hat{R}$ value is greater than 1.1, it indicates non convergence. Thus, the solution is either to reduce $\hat{V}$ the or to increase W, which can be achieved through increasing the iteration number. Further, H. Du et al. suggested to use both PSRF point estimate and the PSRF upper confidence interval (UCI) which is the upper bound of the 95% credible interval to test the convegence [128]. Thus, the MCMC simulation is considered converge if the following conditions are satisfied;

- The PSRF point estimate is close to 1 and no greater than 1.1 [126].
- The PSRF UCI is commonly less than 1.1. The PSRF UCI explains the uncertainity in the PSRF. The uncertainty in the PSRF estimate is influenced by both variances of the chains and the number of iterations. Thus, a higher uncertainty in the PSRF estimate indicates a less accurate convergence assessment, and suggests that increasing the number of iterations is required [128].

The $\hat{R}$ value can be simply calculated by using different statistical packages in R such as coda [129] and rstan [130].

Effective sample size (ESS)

Effective sample size is another quantitative measure diagnostic tool to assess the MCMC convergence. It is a single and dimensionless number that reflects the effec-

tive and independent sample size to measure the efficiency of the MCMC algorithm. It captures the autocorrelation impact of the variability within the chains on the uncertainty in estimates. The ESS can be estimated by using;

$$ESS = N_{\text{eff}} = \frac{N}{1 + 2 \sum_{k=1}^{\infty} ACF_k}$$

where N is the total sample size and AC$F_k$is the ACF value at lag k. The N_{eff}can be greater than N, if the $-1 \overset{\infty}{<} \sum_{k=1} ACF_k < 0$. In addition, the $N = N_{\text{eff}}$, if the $\sum_{k=1}^{\infty} ACF_k = 0$ which indicates each sample provides unique information.

Chapter 3

Statistical analysis of citizen science AI data: A case study of Fall Armyworm invasion in India

3.1 Introduction

Invasive species are one of the key issues threatening the environment and agriculture throughout the world. Plant pest and disease outbreaks seriously affect agricultural economies and threaten their food and nutritional security [131]. A major challenge in these situations is the detection of new pest and diseases by the local communities and scientific researchers. Early detection of incursions can be critical in management practices that have the potential to control the spread of these pest to other areas [132]. A monitoring system of pest and disease risk register is developed in most countries, helping in the detection of a new invasion [133]. However, despite the extensive control measures, new invasions such as Fall Armyworm (FAW) have been reported in India since the middle of 2018 [134, 135].

There is a variety of approaches for detecting pests and diseases, such as qualified inspectors, general public participation (citizen science), remote sensing using satellites and smartphones [74]. Recently, citizen science reporting that depends on assessments by people and farmers has played an important role in detecting pest and diesease occurrence, around the world. Smartphones have become increasingly popular in the past few years providing citizens with access to modern technology, allowing rapid collection of large data sets, including geo-tagged images of pests and their damage to hosts. This

creates an opportunity for introducing novel detection methods based on artificial intelligence (AI) and machine learning (ML) technologies [74]. One such mobile based technological company is Progressive Environmental and Agricultural Technologies (PEAT GmbH), an AgTech startup based in Berlin which has successfully implemented an Android/ Apple based AI/ML application to detect plant damage symptoms, 'Plantix app' [82, 79, 80].

Plantix application has been trained to automatically detect approximately 550 major pests, diseases, and nutrient deficiencies in 35 major crops. A geo-time tagged image is uploaded by a smartphone user to the Plantix app. Then, the app analyses the uploaded image using deep neural networks (DNN) trained on previous data sets. The DNN examines the uploaded image and identifies different health conditions with associated probabilities (similarities)[79, 80, 83]. The DNN selects the health status with the highest probability (top-1 similarity) as the final decision which is then sent to the user with recommendations on the management of the disease [82, 79, 80].

Fall Armyworm (FAW, *Spodoptera frugiperda*) is an invasive insect pest from the Americas, causing massive losses to agriculture in general, and the maize crop in particular, impacting the industry in India since its initial invastion in 2018. Maize (*Zea mays*) is the third-largest crop by area grown in India, spread across an area of 9-9.5 million hectares (mha) with a production of 24–28 million tonnes (mt) [136]. Although, FAW is a polyphagus pest, it has a preference for maize. First reported in Karnataka state in India during May 2018, FAW spread to more than 90% of maize fields across India within 16 months [35, 29]. Mayee et al. [136] reported that FAW caused significant damage leading to a decrease in the maize growing area from 9.2 mha in 2018 to 8.19 mha in 2019. As a result, maize production diminished, leading to a rise in feed prices for poultry, aquaculture and cattle industries, thus impacting outputs from these industries. The Plantix app has been used in India since 2017 [137]. During the 2018-19 seasons, nearly 140,000 data points of maize crop pests and diseases were recorded using the Plantix app.

While Plantix app is successful in diagnosing maize crops in a short time and at the right location, a major challenge is the absence of independent verification. Consequently, there is no established gold standard (reference standard). In this case, it is difficult to determine the false and true diagnoses and to conduct further analysis. Therefore, such mathematical classification methods should be supported by internal

evidence. Hampf et al. [80] used the Plantix app data to detect pests and diseases in maize and soybean crops in the southern Amazon, Brazil, selecting data points with at least a similarity of 0.5 (50%) to ensure reliable diagnoses. Then, the selected data was used to assess the Plantix app accuracy by comparing the diagnostic app results with actual conditions (gold standard). The sensitivity of the Plantix app was reported as 91.51% for maize, although it was lower for soybean diseases and pests. A majority of other studies focused on determining the quality of the app in the classification crop types (names). As an example, Wang et al. [79] measured the app's quality by analysing southeast Indian crop types.

For data classification, latent class analysis or panel diagnosis are the most common methods. The panel diagnosis depends on a group of experts who have the skills required to decide upon the final diagnosis in every case and is based on all the available related case data [138]. The latent class analysis relies on the available diagnostic test results to create a statistical model that establishes a reference standard [139]. Latent class analysis is a type of mixture model that is used for model-based clustering. A Gaussian mixture model (GMM) is the most commonly used around model-based clustering [140].

Recent studies regarding FAW in Indian maize crops have addressed biology, genetic characterization [141, 142, 143], ecology, field survey analysis [144] and controlling and managing the spreading of FAW [145, 35]. According to our knowledge, no study has yet evaluated the temporal spread of infested maize crops and their intensity due to FAW in India. From the Plantix app, a high or low number of FAW notifications cannot solely reflect the intensity of FAW infestation. These notifications may be high because of a significant volume of data collected during the same period. Therefore, comparing FAW reports with all other notifications is essential to understand FAW intensity. This approach can be referred to as *odds*. The term *odds* refers to an estimated rate of the number of occurrence of a certain event (infested maize crops by FAW) to the number of non-occurrences [146]. Hence, this study aims to apply a classification model to distinguish between false and true diagnoses, and to investigate the intensity of FAW infestation on maize crops.

3.2 Materials and methods

3.2.1 Data description

The dataset consists of all notifications regarding maize in India, including those diagnosed as infested with FAW and those without Fall Armyworm infestation. Provided by PEAT GmbH, this dataset comprised the data which was crowd sourced from users of the Plantix app in India. The dataset included various dates from January 1, 2018 until December 31, 2019. Within this period, there were 138359 data points under the 1.1.6 DNN version from 631 districts.

A number of meta-data variables were derived for further analysis (see Table 3.1), including the time stamp, longitude and latitude GPS coordinates and the results of the machine learning as categorical and numerical variables. The categorical variables were saved under the headings "top-1 pathogen name" and "top-2 pathogen name". These two variables define the detection status of each maize diagnosis as either healthy (indicating the maize is in good health), infested with FAW, infested with pathogens, affected by other pests, or suffering from nutrient deficiencies.

The numerical variables were "top-1 similarity" and "top-2 similarity", representing positive integer values ranging from 2 to 100. A large similarity value indicated that the detection was more confident of reflecting the real status of the diagnosed maize. Conversely, similarity values close to one suggest less confidence in detection. It is noteworthy to mention the distinction between top-1 and top-2. The DNN generates several similarity scores with corresponding health conditions for each tested crop where the total sum of these similarities equals 100% (or a probability of 1) . Hence, under the top-1, the highest value is recorded, while top-2 includes the next highest value.

Moreover, this dataset was based on three different deep neural network (Plantix-DNN) versions, where each was applied to a different set of maize cases. For the purpose of our statistical analysis, the research described in this chapter was based on the 1.1.6 version of the app. This is because it is the most recent version, and the analysis was more consistent when it depended on only one specific version. Hence, this version had 134938 data points in 2018-19 which were used in the subsequent analysis. Further, there were 1101 cells in the "District" variable without the district name and therefore labelled with “FAILED”. Accordingly, they were defined manually based on their GPS coordinates using ArcGIS Pro (version 2.8, ESRI, Redmonds, CA, USA).

Table 3.1: Sample data from the Plantix app illustrating the key variables used in statistical analysis.

Date	Longitude	Latitude	Pathogen name		Similarity		DNN	District
			top-1	top-2	top-1	top-2		
01/01/2018	75.73	15.31	Boron Deficiency	Healthy	76	2	1.1.6	Gadag
22/07/2018	75.13	14.77	Spotted Stem-borer	Fall Army-worm	28	18	1.1.6	Haveri
07/12/2019	87.39	25.73	Fall Army-worm	Violet Stem Borer	22	20	1.1.6	FAILED
21/11/2019	81.44	22.02	Fusarium Ear Rot	Healthy	31	6	1.1.15	Jalgaon

3.2.2 Research assumptions

In addition, since the FAW was the primary research focus, we assumed positive results when the Plantix app generated a report for FAW and negative results for all other categories. Further, the second assumption was that the 134,938 maize notifications were classified into three categories:

- FAW: maize crop for which Plantix app reports FAW under the top-1.
- Healthy: maize crop reports as healthy under the top-1.
- Non-FAW: maize crop reports as having other conditions (non FAW pests, diseases, or nutrient deficiencies).

Figure 3.1 displays the 134,938 maize notifications on the map of India. It can be seen that the red points (FAW notifications) are more abundant in the South compared to the North, while the green (Healthy) and the blue (Non-FAW) dots are more frequent in the North. Therefore, the third assumption was that the data was divided the notifications into two zones, in order to see the difference in the spread and where the infestation started during the period of analysis. The fourth assumption, this study mainly depended on top-1 variables rather than the top-2 as it was the most accurate when diagnosing variables using the Plantix app.

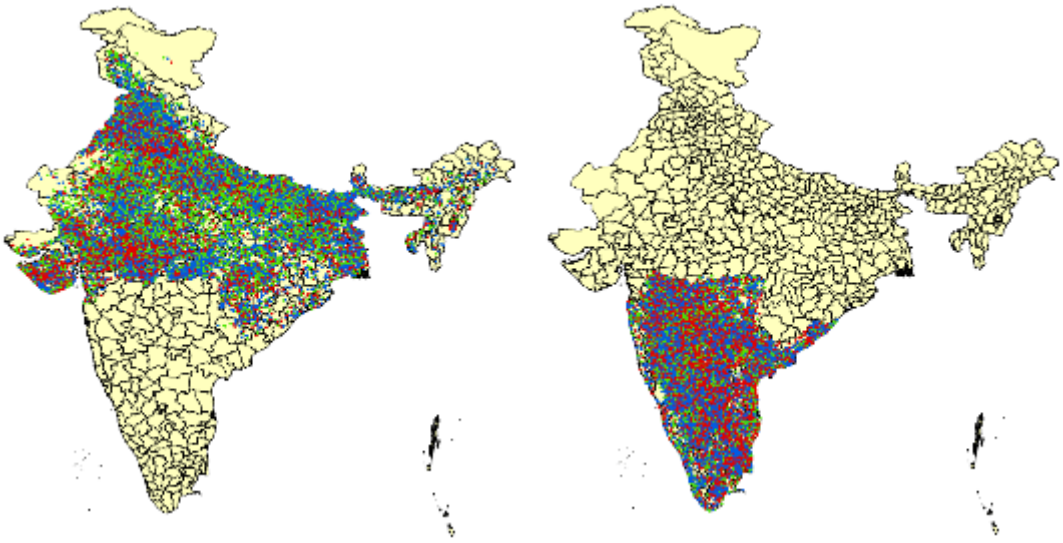


Figure 3.1: The Plantix app top-1 notifications for the maize crops in the South (left) and North (right) of India during the invasion period (2018-2019) with FAW (red), Healthy (green) and Non-FAW (blue).

3.3 Classification model

A classification model was applied to assign each tested maize crop into a True (T) or False (F) case, with True (T) associated with accurate diagnoses and False (F) with the app failing to diagnose the maize crop correctly. The classification process was based on the "top-1 similarity" variable. Therefore, the True class includes high similarity values, while the false class includes low similarity values. Moreover, the classification model was utilised separately for the three categories FAW, Healthy and Non-FAW.

The classification model was based on a mixture of two normal distributions with the expectation-maximisation (EM) algorithm, which was used to estimate the parameters. Although the similarity values were discrete numbers ranging from 2 to 100, the normal distribution was chosen as a first approximation among possible continuous distributions due to its flexibility and widespread applicability in statistical modeling. Furthermore, the normal distribution is commonly employed in latent class models because of its mathematical properties, which facilitate estimation and classification. Further, with a large sample size of over 10,000 for each category, the normal distribution provides a robust starting point for modeling the data.

The mixture of two normal distributions represents a sum of two normal distributions with different weights. The expectation-maximisation (EM) algorithm is a statistical

method to estimate the parameters for the mixtures[109, 110, 111, 108, 107]. Opsteegh et al. [94] applied this algorithm successfully to classify the results of the RIVM ELISA test to detect the Toxoplasma gondii parasite in sheep into healthy (Negative) and unhealthy (Positive) groups. In this study, we implement the algorithm three times, once for each category independently: FAW, Healthy, and Non-FAW. For each category, the algorithm is applied to identify True and False groups, resulting in three separate sets of classifications. This represents a novel application to this type of data. Additionally, while most classification methods typically distinguish between positive and negative cases, our work specifically focuses on classifying data into True and False categories. This unique approach offers a new perspective on analysing the data.

In this context, the mixture model was appropriate as the presented data included both the latent and observed variables, where both variables are to explain the data distribution. The latent variable here was a binary variable, with two classes representing False and True classifications. The observed variable was the "top-1 similarity" variable. Hence, the model can define the heterogeneity between the observed data points by estimating the probability of each data point belonging to either the False or True class.

Therefore, the mixture model[109, 110, 111, 108, 107] for each category can be expressed in the following way:

$$\mathcal{N}\left(\mu_F, \mu_T, \sigma_F^2, \sigma_T^2, \omega\right) = (1-\omega)\,\mathcal{N}_F\left(X; \mu_F, \sigma_F^2\right) + \omega \mathcal{N}_T\left(X; \mu_T, \sigma_T^2\right) \tag{3.1}$$

where $0 \leqslant \omega \leqslant 1$ is the mixture proportion, representing the proportion of the tested maize crops that belonged to the True class. Thus, $(1-\omega)$ indicates the proportion of False readings. The one dimensional random variable X is the "top-1 similarity" score, where each score is assumed to be independent. This is because each image is diagnosied separately by the Plantix app, where the diagnosis outcome for one crop does not impact the diagnosis of another. While it is possible for nearby crops to have similar symptoms, which could result in indirectly correlated scores, the independence assumption remains reasonable for the purposes of this analysis. Further, $\mathcal{N}_F\left(X; \mu_F, \sigma_F^2\right)$ and $\mathcal{N}_T\left(X; \mu_T, \sigma_T^2\right)$ are univariate Normal distributions. Thus, in this model, there are five parameters $\mu_F, \mu_T, \sigma_F^2, \sigma_T^2$ and ω which need to be estimated.

3.3.1 Parameter estimation

Accordingly, the five parameters $\mu_F, \mu_T, \sigma_F^2, \sigma_T^2$ and ω were estimated by the EM algorithm. The EM algorithm can be applied iteratively to cycle between E (Expectation step) and M (Maximisation step) steps starting from the initial values of each parameter. The E-step involves calculating the expected value of the log-likelihood function, given the current estimates of the parameters and the observed data. The M-step then updates the parameter estimates by maximising the expected log-likelihood obtained in the E-step, hence, improving the model's fit to the data.

The initial values of μ_F, μ_T, σ_F^2 and σ_T^2 in this study were estimated using k-means clustering where $k = 2$, where k-means clustering is an algorithm that partitions data into k clusters based on minimising the distance between data points and their respective cluster centers. The initial value of ω was computed using the mean (weight) of one cluster. After that, the E step assigns the probabilities of each data point in X belonging to the True class and False class are calculated based on current parameter estimates values through the responsibility function $\widehat{\gamma}_i$. The probability $\widehat{\gamma}_i$ of data point i in X belonging to the True class given the current estimates of the model parameters $\mu_F, \mu_T, \sigma_F^2, \sigma_T^2$ and ω is represented as

$$\widehat{\gamma}_i = \frac{\omega \mathcal{N}_T\left(X = x_i; \mu_T, \sigma_T^2\right)}{(1-\omega)\,\mathcal{N}_F\left(X = x_i; \mu_F, \sigma_F^2\right) + \omega \mathcal{N}_T\left(X = x_i; \mu_T, \sigma_T^2\right)} \tag{3.2}$$

The complement of $\widehat{\gamma}_i$ is $(1 - \widehat{\gamma}_i)$, and it infers the probability that each data point belongs to the False class. The normalisation term in the denominator guarantees that the sum of probabilities for a single data point is equal to one. On the other hand, the M-step updates the parameter values based on $\widehat{\gamma}_i$ in the E-step. Hence, it allows us to find the parameters' values that maximise the likelihood function of the observed data [109, 110, 111, 108, 107]

. Further, this iterative process between the E and M steps was repeated until convergence, with the convergence threshold set as 1×10^{-6}. The EM algorithm was run in R 4.2.1 and tested using a normalmixEM function in mixtools package [147].

Further, to assess the goodness of fit of the mixture model, the Q-Q plot and a Kolmogorov-Smirnov (K-S) test with a parametric bootstrap method was implemented. The parametric bootstrap approach was used to correct for inaccuracies in the critical

values of the K-S test, which were caused by estimating model parameters from the data [148]. In this process, 1,000 bootstrap samples were were generated from the fitted mixture model. Then, the K-S statistic for each sample was computed in oreder to form a bootstrap distribution, which facilitated to determine an adjusted p-value by comparing it to the observed K-S statistic [148].

3.3.2 Weekly temporal distribution of False (F) and True (T) groups in North and South zones

After fitting the mixture model of two normal distributions with latent classes (True and False) using the whole dataset for each category, a temporal distribution for each class was explored, in order to estimate the weekly expected notifications for both classes. The temporal distributions were built for each category (FAW, Healthy, Non-FAW), in North and South zones. Note, the subsequent actions are only explained for the FAW category in the North zone.

At first, the daily notifications were converted into weekly notifications, resulting in a total of 104 weeks for the whole two years (2018-19). This transformation was performed to simplify the analysis without impacting its accuracy as well as the fact that there was not enough data for 720 days. Following that, for each week w_j $(1 \leq j \leq 104)$, the number of notifications $(frequency)_{ij}$ at each similarity rate $(2 \leq i \leq 100)$ was determined from the data to estimate the weekly expected notifications of F class w_j^F and T class w_j^T. These expectations are given by:

$$E\left[w_j^T \mid FAW_{North}\right] = \sum_{i=2}^{100} (frequency)_{ij} . \widehat{\gamma}_i \tag{3.3}$$

$$E\left[w_j^F \mid FAW_{North}\right] = \sum_{i=2}^{100} (frequency)_{ij} . (1 - \widehat{\gamma}_i) \tag{3.4}$$

where $E\left[w_j^T \mid FAW_{North}\right]$ and $E\left[w_j^F \mid FAW_{North}\right]$ are the weekly expected value of FAW in the North zone in T and F classes, respectively using the final parameters estimated from the converged EM algorithm of the mixture model of FAW category and responsibility function (eq. 3.2). The $E\left[w_j^T \mid FAW_{North}\right]$ reflects the weekly

expected number under the FAW category in the North zone that the Plantix app successfully diagnosed, while the $E\left[w_j^F \mid FAW_{North}\right]$ of FAW representes the weekly expected number in the North zone that the Plantix app failed to diagnose correctly. Here, $frequency_{ij}$ represents the number of observations within the FAW category at i similarity rate and in a week j.

This method was applied for all categories (FAW, Healthy and Non-FAW) and in both spatial zones. As stated described that

FAW is assumed as the positive category, while Healthy and Non-FAW are assumed as negative categories. Hence, the $E\left[w_j^T \mid FAW_{zone}\right]$ representes the weekly expected number of True Positive (TP). Similarly, $E\left[w_j^T \mid \text{Healthy}_{zone}\right]$ and $E\left[w_j^T \mid \text{Non-FAW}_{zone}\right]$ denote the weekly expected numbers of True Negative (TN) for Healthy and Non-FAW categories, respectively. In addition, the $E\left[w_j^F \mid FAW_{zone}\right]$ indicates the weekly expected number of False Positive (FP), while $E\left[w_j^F \mid \text{Healthy}_{zone}\right]$ and $E\left[w_j^F \mid \text{Non-FAW}_{zone}\right]$ represent the weekly expected number of True Negative (TN) for Healthy and Non-FAW categories, respectively. Table 3.2 summaries these terminologies.

Table 3.2: Four possible outcomes of the weekly notifications can be estimated based on the classification models, which is based on "top-1 similarity" and three categories (FAW, Healthy and Non-FAW) under the "top-1 Pathogen name" variable.

		Plantix app variable "top-1 Pathogen name"	
		Infested with FAW (Positive)	Not infested with FAW (Negative)
Classification model	True (T)	True positive (TP)	True Negative (TN)
	False (F)	False positive (FP)	False Negative (FN)

3.3.3 Intensity of weekly FAW infestation

In order to estimate the extent of the FAW infestation, the number of TP notifications needs to be interpreted in the light of the overall reporting intensity. Thus, the large number of TP notifications can be associated either with high intensity and medium (or even low) reporting rate, or low intensity but a very large number of reports. The solution is to scale TP by TN, resulting in odds. Thus, the *odds* formula of infested maize crops by FAW is:

$$odds = \frac{TP}{TN} \tag{3.5}$$

where the *odds* ranges between 0 and infinity. We considered two scenarios for TN, where TN either represents TN cases of Healthy or TN represents TN cases of both Healthy and Non-FAW. If $odds > 1$, the maize crops infested by FAW outweigh the number of non-infested by FAW, and if $odds < 1$, the number of non-infested by FAW is higher. Consequently, when $odds = 1$, the number of maize crops infested by FAW and non-infested by FAW are equal [149]. Additionally, the *odds* can be converted to probability:

$$probability = \frac{odds}{odds + 1} \tag{3.6}$$

where this probability indicates the likelihood of maize crops being infested by FAW over week. For instance, if the probability is 0.6, it means that 60% chance of maize infestation with FAW in that week.

3.4 Results

3.4.1 Descriptive analysis and data visualisation

Bimodal distribution of the similarity index

The similarity rates for the three categories (FAW, Healthy, Non-FAW) in the whole data set, as well as for the North and the South of the country, are shown in Figure 3.2. A considerable number of notifications (63.4 %) have a low similarity rate (<50), while a few notifications (0.72 %) have similarity rates greater than 90. Compared to North India, South India recorded 20.3 % more notifications, despite the North's area being larger than the South (see Figure 3.1), thus, reflecting a higher proportion of maize growing land in the South.

Further, two peaks can be observed in each category in both North and South India, and more notifications are present in the left peak, with the exception of the Healthy category. We use this property of each category of notifications, to classify False (F) and True (T) classes in the mixture model, with F and T classes as a latent variable. Therefore, in this situation the mixture of two normal distributions can be used as a classification model, in order to predict the four possible outcomes of a diagnostic test: true positive (TP), false positive (FP), true negative (TN), and false-negative (FN).

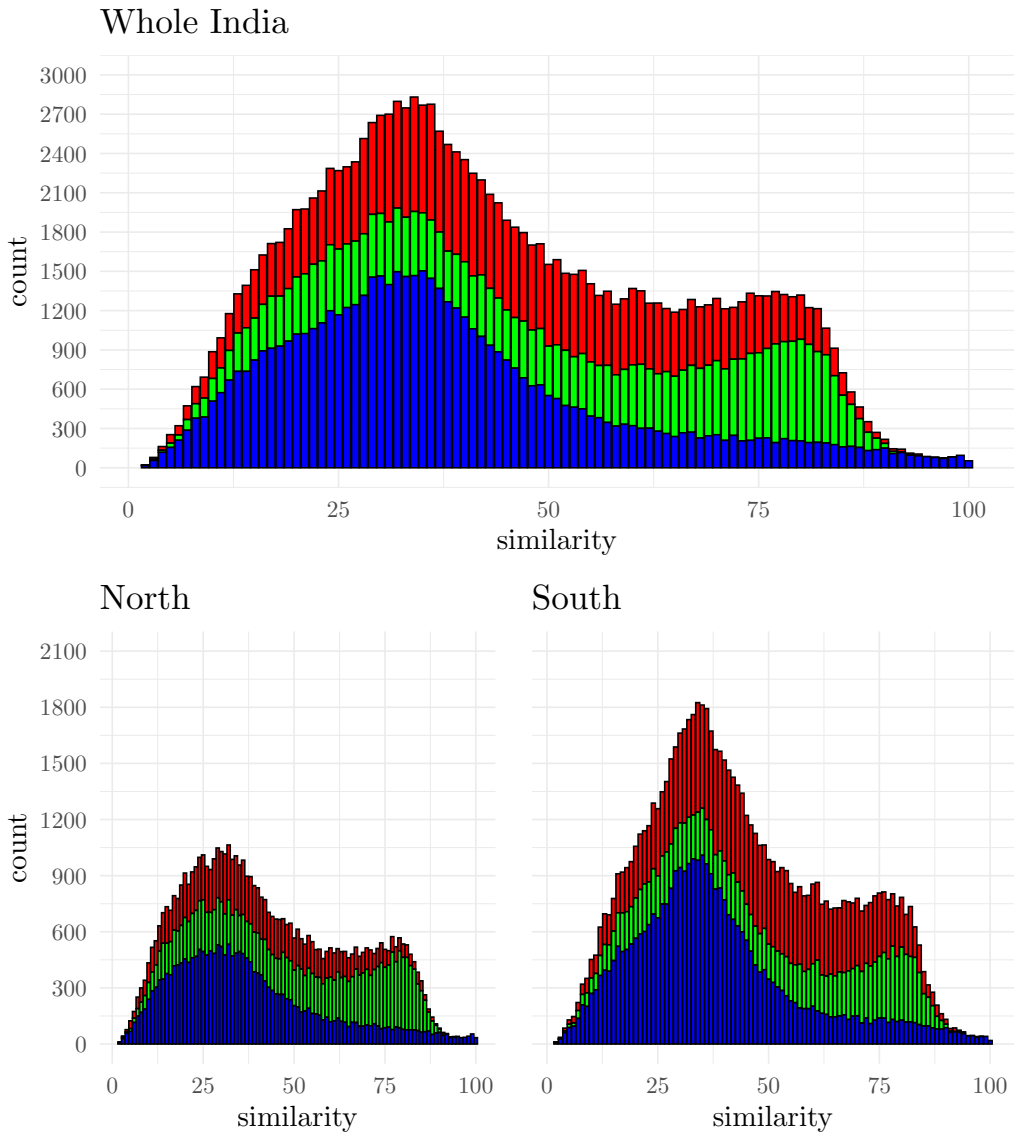


Figure 3.2: Stacked histograms of similarity rate (2-100) for maize crop in pan India during 2018-19 (n=134938), for North India (n=53779), and South India (n= 81159), showing FAW (red), Healthy (green) and Non-FAW (blue).

Temporal notifications

Figure 3.3 shows the weekly notifications in both zones. Kharif (June - October) and Rabi (October - February) are the typical seasons for growing maize in India. Although the dates vary from year to year, and between locations, we assign the dates of 3rd June and 9th October as the start and end of the Kharif season, whereas we consider the dates 10th October and 28th February to be when the Rabi season begins and ends.The South zone notifications under each category outnumbered the North's notifications over the first 10 weeks. Further, the FAW notifications in the South were higher when compared

to the notifications from the North for the entire period, with the exception of the time between March 2018 and the beginning of Kharif 2018. During North Kharif seasons, the notifications in the Healthy category surpassed the onse in the South, whilst the converse was true during the Rabi Seasons. Furthermore, the South accounted for more Non-FAW notifications than the North during March and mid-July of 2018.

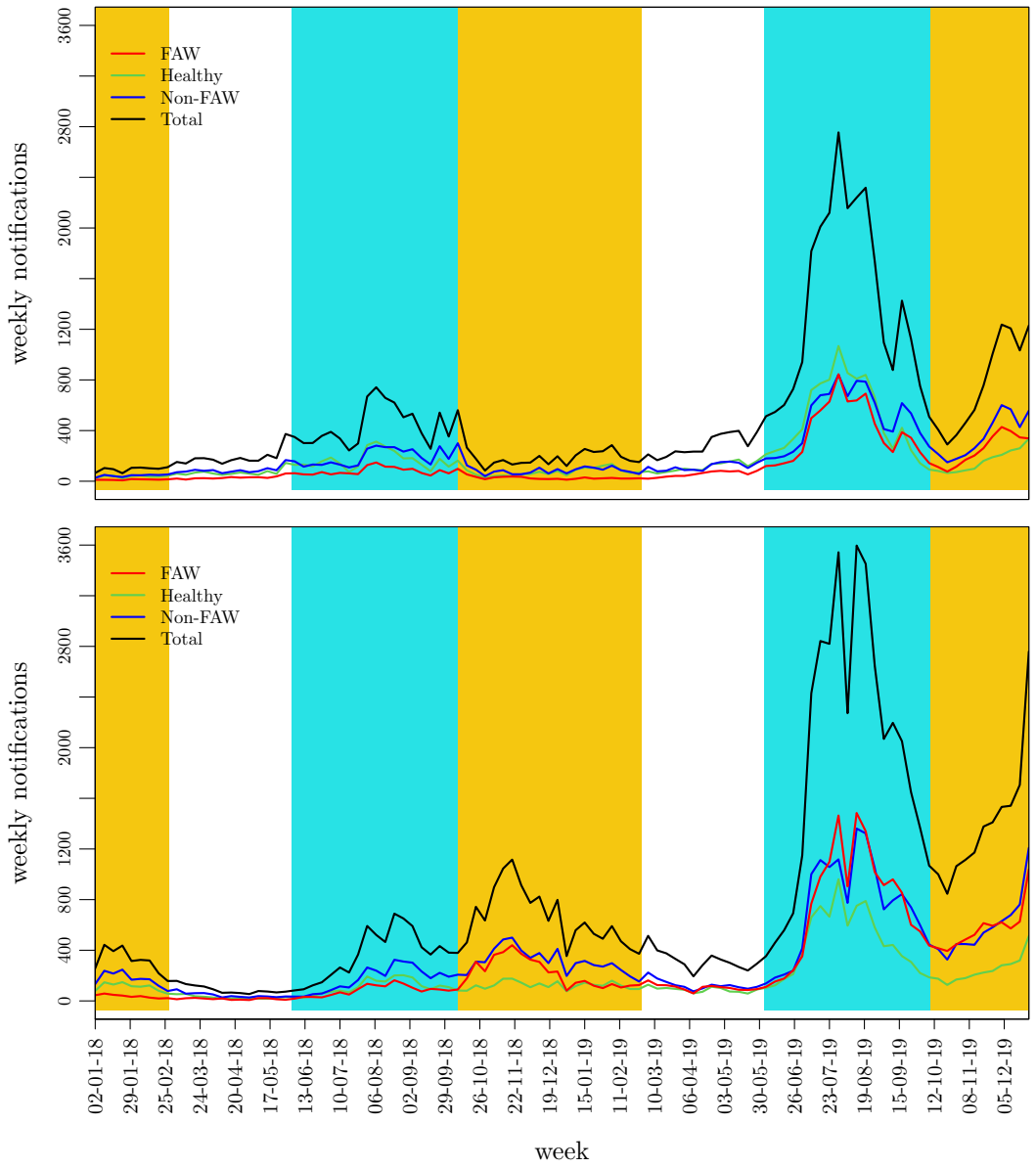


Figure 3.3: The weekly number of maize crops diagnosed by the Plantix app in the North (top) and South (bottom). The red curve is the True class, and the blue curve is the False class. The light blue area is Kharif season (3rd June to 9th October), the orange area is Rabi maize season (10th October to 28th February).

The results of the weekly notifications show that most of the notifications in both years have been in Kharif seasons, except for the FAW category in 2018, which had more

Rabi than Kharif notifications. This can be interpreted by comparing Figure 3.3 with Figure 3.4, where the number of users of Plantix app were more active during Kharif and Rabi seasons. Especially during Kharif, the number of users was higher than at other times as it is a major monsoon season to grow crops. In addition, there were more users in 2019 than in 2018. This might mean that the number of Plantix app users was increasing or they were becoming more experienced at using the app.

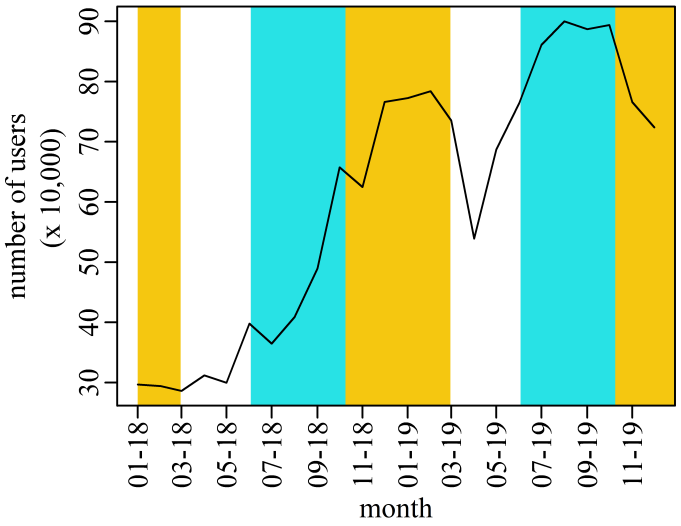
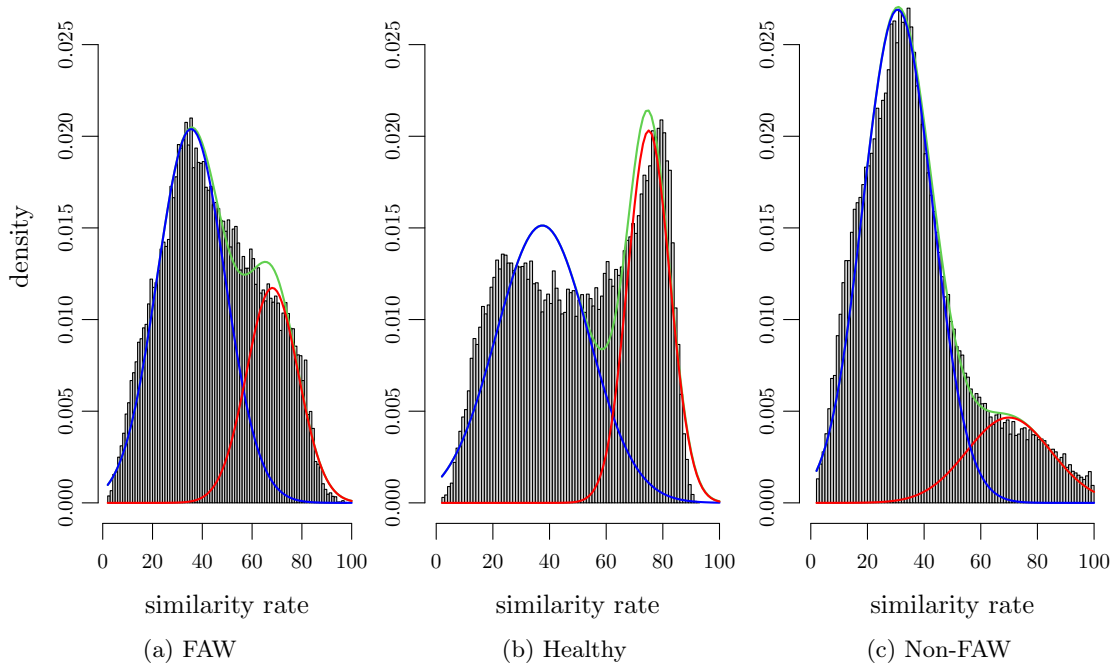


Figure 3.4: Monthly total number of the Plantix app users in the whole of India within the study period (2018-2019).

3.4.2 Classification model: mixture of two normal distributions



(a) FAW (b) Healthy (c) Non-FAW

Figure 3.5: Fitted two univariate normal distribution mixture model (green curve) based on similarity rate along with probability fitted models for the false group (blue curve) and true group (red curve). The intersection point between the blue and red curve is the cut off between two groups. The cut off value is 57 for FAW, 63 for Healthy, and 57 for Non-FAW

The mixture model is used to filter out the low-similarity notifications, interpreted here as False (F) in contrast to high similarity True (T) cases. The histograms from Figure 3.2 and the corresponding fitted curves using Eq 3.1, are shown in Figure 3.5. The estimated parameters of the mixture models are presented in Table 3.3.

Table 3.3: Parameter estimations based on mixture of two univariate normal distribution with the EM algorithm.

Category	False group parameters	True group parameters
FAW	$\omega = .70$, $\mu = 35.51$, $\sigma = 13.60$	$\omega = .30$, $\mu = 68.13$, $\sigma = 10.37$
Healthy	$\omega = .62$, $\mu = 37.48$, $\sigma = 16.34$	$\omega = .38$, $\mu = 75.02$, $\sigma = 7.46$
Non-FAW	$\omega = .83$, $\mu = 30.68$, $\sigma = 12.25$	$\omega = .17$, $\mu = 69.92$, $\sigma = 14.84$

The visual inspection of the Q-Q plots for each category in Figure 3.6 suggests a good fit for the mixture models. This is because across all three plots there is an indication of linearity, where there are strong linear relationships between the points of empirical CDF

and theoretical CDF. Further, the points do not consistently deviate above or below the line, indicating that the theoretical distribution is capturing the data's behavior well. In addition, the calculation of the Kolmogorov-Smirnov (K-S) test with a 1000 parametric bootstrap method was implemented for each mixture model. The p-value for each fitted mixture mode was above 0.05, indicating that there is insufficient evidence to reject the null hypothesis. Hence, we can conclude that each mixture model is a reasonable fit for the data.

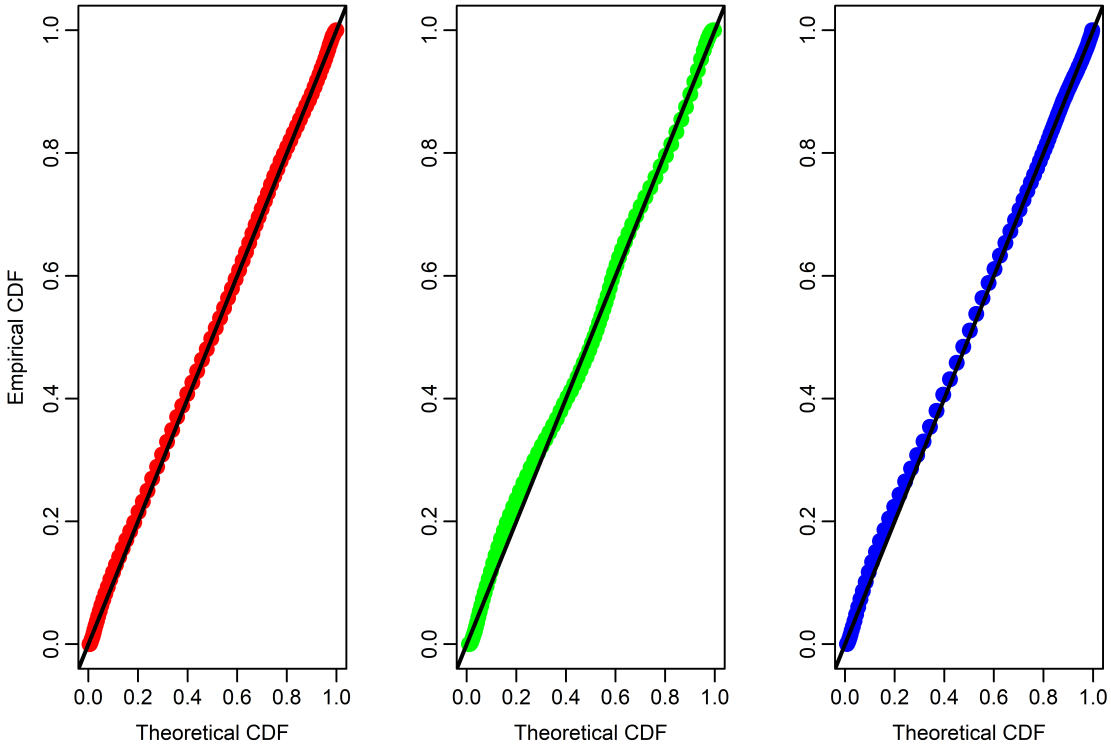
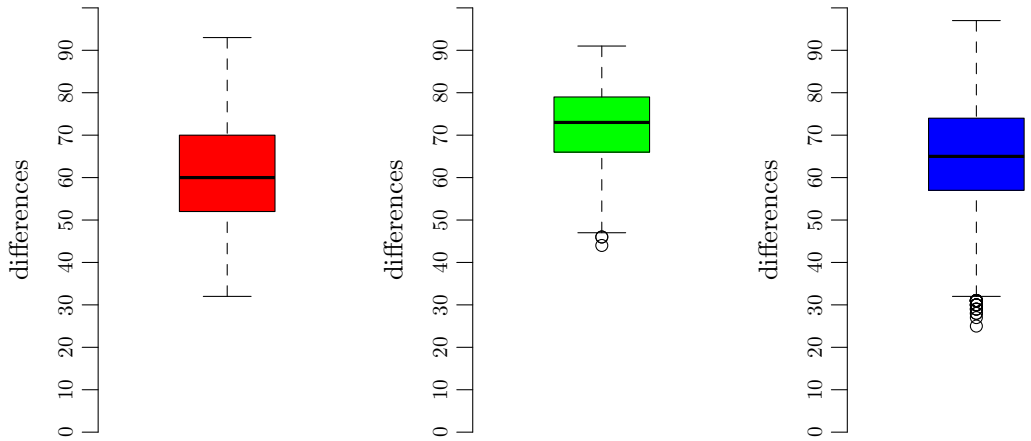


Figure 3.6: Goodness of fit for mixture model using Q-Q plot.

One more analysis which reflects the goodness of fit is comparing the fitted mixture model with observation data for top-1 similarity" and corresponding rates of "top-2 similarity". It is notable that cut-off points between F and T groups were at a similarity value of 57 for FAW, 63 for Healthy, and 57 for Non-FAW (see Figure 3.5). For the true groups, we considered them as the starting rate. To check the accuracy of the fitted model, these cut off rates were used to find differences between True rates of "top-1 similarity" and corresponding rates of "top-2 similarity". The differences were consistently above 30 as can be seen in the boxplots (Figure 3.7). Hence, the results show that the presence of such gaps indicates that the fitted model was successful in

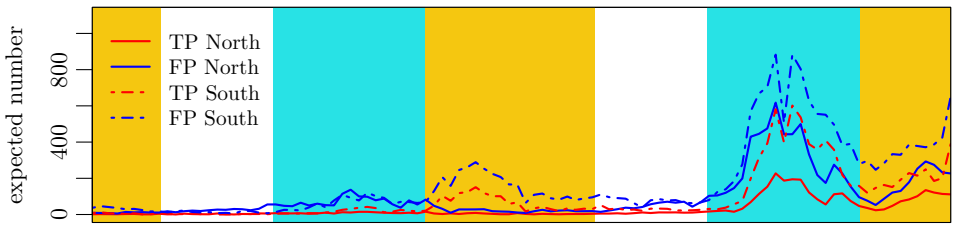
determining T and F groups. In other words, a significant difference between True rates of "top-1 similarity" rates and corresponding rates of "top-1 similarity" indicates that the model was successful in accurately predicting the true diagnosis cases by using the Plantix app. In contrast, if the distinguished value is small, such as 10 or less, it reflects that the classification model may struggle to make a clear determination.



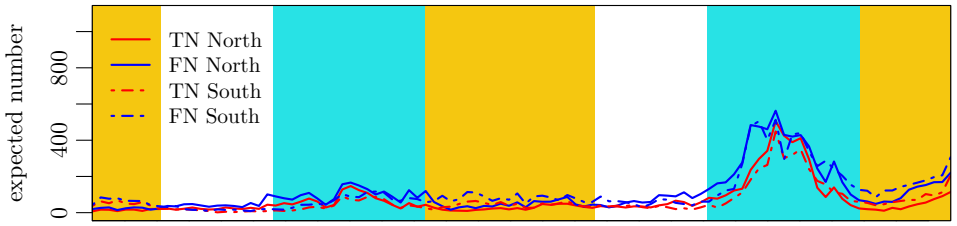
(a) FAW (b) Healthy (c) Non-FAW

Figure 3.7: Boxplots of similarity rates difference between “top-1 similarity” and “top-2 similarity”, where the top-1 started at the cut-off rate with (a) the FAW at 57, (b) the Healthy at 63 and (c) the Non-FAW at 63.

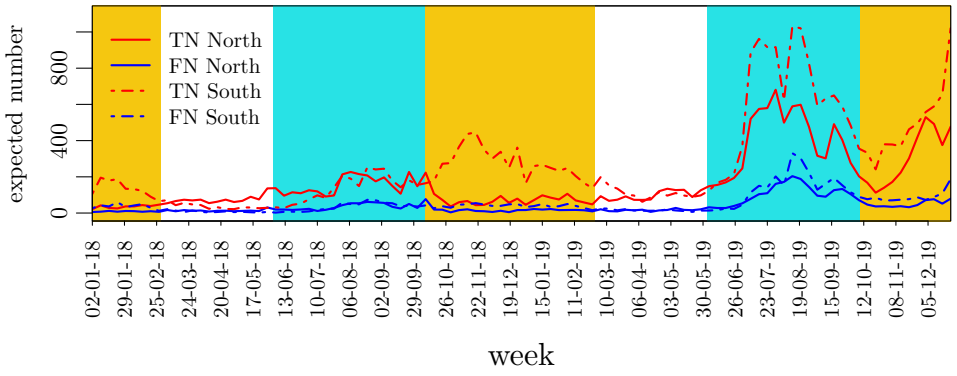
3.4.3 Weekly temporal distribution of F and T groups in North and South zones



(a) Weekly expected number of FAW cases based on mixture normal distribution of FAW.



(b) Weekly expected number of Healthy cases based on mixture normal distribution of Healthy.



(c) Weekly expected number of Non-FAW cases based on mixture normal distribution of Non-FAW.

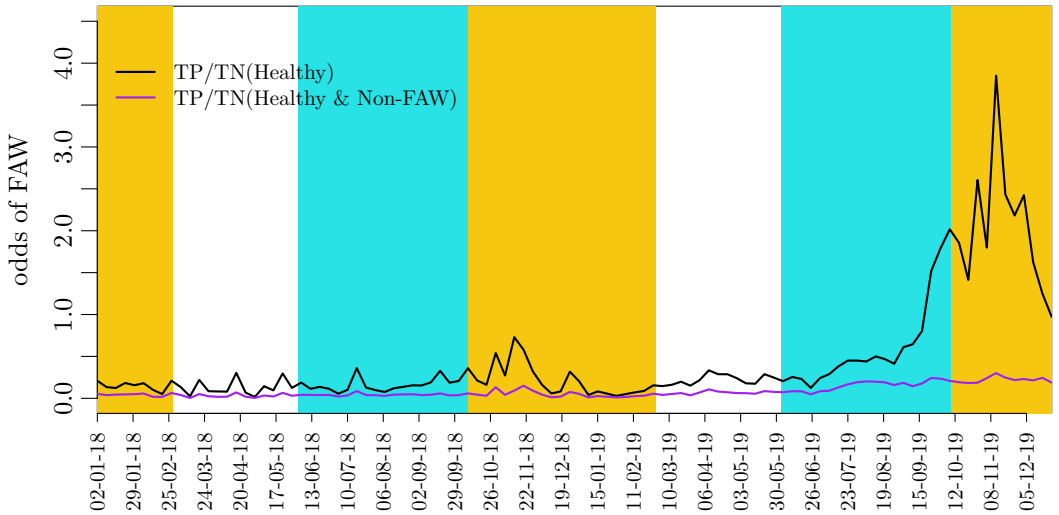
Figure 3.8: Weekly expected number of infested and non-infested maize crops in 2018 and 2019 in North (solid line) and South (dashed line) India, where the red curve is for the True groups, the blue curve is for the False groups, and the light blue area is the Kharif season (3^{rd} June to 9^{th} October) and the orange area is the Rabi season (10^{th} October to 28^{th} February).

The weekly expected notifications of tested maize crops under each class (F and T), and by category, was estimated by equations 3.3 and 3.4, using the estimated parameters of the mixture models (see Table 3.3). The expected weekly notifications of both T and F classes for the three categories FAW, Healthy and Non-FAW during 2018-19 can be seen in Figure 3.8. Figure 3.8 shows that the weekly notifications are distributed quite

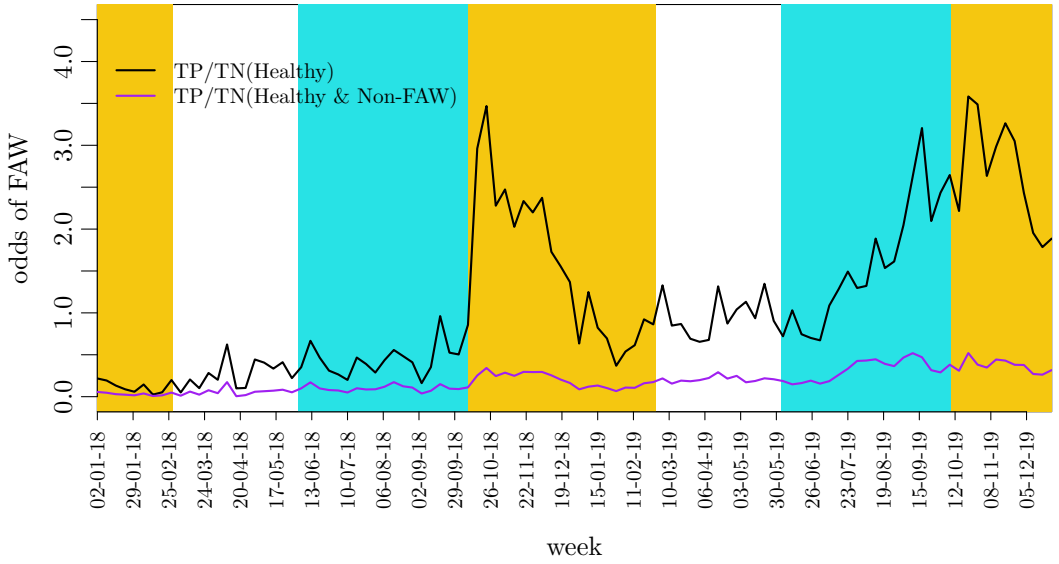
differently between the South and the North of India. For North India, the notification levels are at the lowest during the first weeks of both years while they peak during the Kharif season. A significant increase in notifications was observed during Rabi 2019. In South India, there were notifications in the early weeks of 2018, peaking during the Rabi season in 2018, for both FAW and Non-FAW. The notifications peaked during the 2019 Kharif season for all the three categories (FAW, Healthy, Non-FAW), and again in Rabi 2019. In both regions, the number of notifications in 2019 was higher than in 2018. The False groups outnumbered True groups, suggesting that overall similarity tends to be low.

3.4.4 The odds of FAW

Figure 3.9a shows the intensity of FAW in the North zone and Figure 3.9b for the South, using the *odds* formula (3.5). By comparing these two Figures with Figure 3.8, Figure 3.8 shows that the highest peak of FAW in both zones occurred during Kharif 2019. However, both Figures 3.9a and 3.9b show different patterns. Figure 3.9b illustrates that the FAW infestation reached its peaks during both Rabi 2018 and Rabi 2019 seasons. The intensity of FAW in the North zone shows that the highest peak of FAW occurred in Rabi 2019 (see Figure 3.9a). Therefore, relying solely on the number of TP notifications can result in incorrect conclusions when determining the FAW outbreak behaviour over time. However, by integrating the True plots (TP, TN), the intensity of FAW in maize fields can be identified.



(a) North India.



(b) South India.

Figure 3.9: Weekly FAW intensity in the maize fields over the whole period (2018-19). The plots were constructed by using the *odds* formula. The black curve represents the intensity of FAW in maize fields when TN cases are only from truly Healthy maize notifications. The purple curve represents the intensity of FAW in maize fields when TN cases are from both truly Healthy and Non-FAW maize notifications.

There were differences between North and South India in the FAW invasion timing and level of occurrence in the maize fields based on TP for FAW and TN for the Healthy category (see the black curve in both Figure 3.9a and 3.9b). In South India, the FAW invasion started at the beginning of Rabi 2018. Moreover, there was a noticeable FAW infestation during 2019, there was a significant increase from the beginning of the Kharif

season to a peak in the middle of the Rabi season with 78% (using formula 3.6, black curve). It is worth noting that in South India, the FAW invasion peaked during both 2018 and 2019 Rabi seasons, and they were at approximately the same level in both the years. In North India, the FAW outbreak began at the end of Kharif 2019, and it reached a peak in the Rabi season with approximately 79% infestation (using formula 3.6, black curve). Moreover, the FAW outbreak was more prevalent in the South than in the North of India. On the other hand, the intensity of FAW in both zones when TN cases are from both truly Healthy and Non-FAW maize notifications (see the purple curve in both Figure 3.9a and 3.9b) show almost flat curves, indicating that there are very few instances of FAW. This pattern raises concerns that there may be issues with the data, which the Chapter 4 explains.

3.5 Discussion

Citizen science apps such as Plantix app are very useful for reporting pest infestations, allowing threats to be detected rapidly, with high temporal resolution, and with a wide spatial coverage of pest distribution. For statistical analysis, it is essential that both positive and negative notifications are included, and mobile apps such as the Plantix app can provide both. This study provides a systematic method for distinguishing between True and False notifications, as well as estimating the intensity of infestation in various species.

The distribution of similarity index is bimodal in all cases, and is particularly pronounced when the notifications are identified as Healthy. This allows to researchers to interpret the two peaks as representing True and False readings, for FAW ("positive") and Healthy or Non-FAW ("negative") crops, respectively using mixture of two normal distributions mixture models. Opsteegh et al. [94] demonstrated that the two normal distributions mixture model can be used to classify diagnostic test results into positive and negative groups when gold standards are unavailable. The difference between our results and this previous study is the objective of the classification. In our approach, we used the classification to differentiative between T and F while positive and negative groups were defined.

Moreover, Hampf et al. [80] considered only images with a top-1 similarity of at least 50 to be as true diagnosis. The method used in this study is superior to the one applied by Hampf et al. [80], as it uses the evidence from the actual data rather than an

arbitrary threshold value of 50. This discrepancy is because they did not depend on a statistical classification framework. It may be that a similarity of less than 50 indicates a high level of confidence in true diagnosis, thus using a cut-off value of 50 may not always represent a good hypothesis.

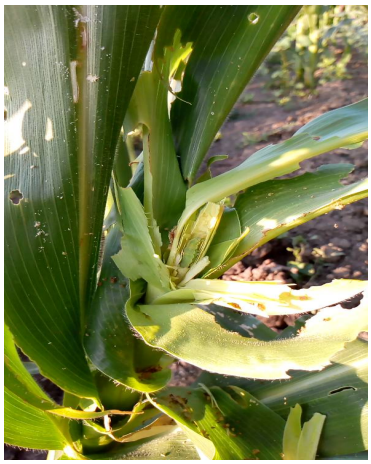
FAW in regions and seasons

A comparison of the *odds* calculation based on the model to the time series based on raw readings showed that the actual progress only occurred in the growing seasons (Kharif and Rabi). Moreover, the FAW infestation was spreading more in Rabi seasons than in Kharif seasons, and more in 2019 than 2018. This finding is in agreement with the estimation that the FAW infested 3.76% of the actual sown area of maize in Kharif 2018, while during Rabi 2018, it infested 27.9% [30]. One plausible explanation for this result might be that the rainfall in Kharif is higher than in the Rabi season due to the monsoon period, where the heavy and light rainfall washes out the first and second instar of larvae from the plant and kill a significant number of them [150]. The *odds* calculation also allowed us to distinguish between the South, where the invasion was earlier, and the North, where it was delayed. This findings agree with Suby et al. [29] who demonstrated, using offical notifications, that FAW started in South India and moved towards North.

FAWs invaded maize in early 2018

In this study, the *odds* calculation showed that the invasion started earlier in 2018, before it was officially documented (mid of 2018) [134, 135]. This finding is in agreement with Figure 3.10 which cites evidence of FAW availability on 21/01/2018 and 23/01/2018. These figures were extracted from the Plantix app server, and their similarity rate is greater than 80. Therefore, the results are inconsistent with the FAW's official detection time of FAW in India, which was in June 2018 [30]. This result emphasises the importance of early data analysis in pest invasion management, and that collecting data without analysing it is not enough. The Plantix app could have been the first to report the invasion in India and this remains a lost opportunity. We suggest that applications such as this should prepare a risk register to demarcate areas by geo-tagging pest and diseases at the country level. This will directly benefit bio-risk assessment and management on invasive pest and diseases.

(a) FAW recorded in 10/01/2018 in Karnataka state, South zone.



(b) FAW recorded in 21/01/2018 in Gujarat state in North zone

(c) FAW recorded in 23/01/2018 in Gujarat state in North zone

Figure 3.10: Examples of photos from the Planxtix app of FAW larvae or damage caused by FAW recorded in January 2018.

Advantages and challenges of citizen data

A unique feature of the Plantix app is the ability to keep track of healthy (negative) and unhealthy (positive) notifications. This feature allows us to estimate the level of intensity of pest or disease in a given geographical area. Therefore, citizen scientists should be encouraged to collect both healthy and diseased data in order to increase the accuracy of detection. Availability of accurate GPS coordinates in the dataset collectes by mobile application, allows estimation of the FAW spread across North and South India. One challenge is that some users might submit photos from places other than their crop fields (such as their homes or cafes) [137]. A large proportion of notifications associated with FAW or other categories was characterised by relatively low similarity.

It is possible that the algorithm has difficulties distinguishing between symptoms caused by different pests (e.g. FAW and others such as Velvet Stem Borer). By combining the detection using mobile app with statistical analysis using the mixture model and odds calculation, we were able to reconstruct the progress of the FAW invasion in India within the study period (18-2019) (Figure 3.9a & 3.9b).

Chapter 4

Diagnostic accuracy of mobile application: A Bayesian meta-analysis approach

4.1 Introduction

The Plantix app, an AI diagnostic application, has become widely popular among farmers and has enhanced the database for researchers [68]. However, some concerns were raised about its accuracy in diagnosing crop problems. For example, Siddiqua et al. [68] concluded that while the Plantix app users feedback indicates that the app is user-friendly, other users comments suggested that the app diagnosis is imperfect as it cannot correctly identify crop problems. In addition, in the previous chapter of this work, the fitted mixture model of the FAW category indicated that the Plantix app has limitations in identifying FAW in maize crops, as shown by the many observations that fall below the false positive curve. Therefore, this inaccuracy can cause concerns among the app users, who are mainly farmers, and the app developer. As a result, the assessment of diagnostic test accuracy is important to quantify this accuracy and reassure all stakeholders.

However, to evaluate the accuracy of the app, a challenge arises. The Plantix app gold standard database for comparison is absent. Hence, the Plantix app accuracy becomes difficult to determine. Moreover, to the best of our knowledge, there is no literature available regarding the app's accuracy when identifying any of the over 600 plant damages among 30 different major crops. To address this challenge, it can be

useful to narrow the focus into one single plant damage source in a solitary crop.

Therefore, the previous chapter can be considered as the starting point to estimate the sensitivity and specificity for the app's accuracy when detecting FAW in the maize crop. Thus, the mixture models results rely on point estimates derived using the EM algorithm in the previous chapter, can be used to build a 2×2 diagnostic table for the entire country of India. This is because the mixture models can be used to calculate the probability of correctly detecting FAW, it is known as sensitivity, and the probability of correctly identifying the absence of FAW, which is referred to as specificity [151]. However, to have a more precise estimation, it is necessary to take into account that the India covers a vast area and contain factors relate to the app and the user.

India has an area of around three million square kilometers [152]. In addition, each Indian state has an independent government, and most differ in terms of languages, environmental factors and citizen scientists skills. For example, India has more than 1600 native languages and 22 official languages throughout its states [153]. However, the Plantix app currently supports eight Indian languages [74]. As a result, it can cause difficulty for some farmers who are not familiar with other accessible languages. Moreover, user skills in taking images and variations in infestation levels can cause variability in both sensitivity and specificity, where there is less diversity in these factors within the state.

All these factors support a state-specific approach to account for regional differences instead of providing a single result for all of India from Chapter 3. Hence, the pooled estimates of sensitivity and specificity across multiple states should be considered to evaluate the diagnostic test for a particular pest. The pooling is performed through meta-analysis of diagnostic test accuracy [154]. This approach can significantly impact the accuracy of the estimation of Se and Sp, and provides a more comprehensive evaluation. Comprehensive literature review of the meta-analysis of the diagnostic test accuracy is provided in the following subsection.

4.1.1 Review of meta-analytic models for diagnostic test studies

Meta-analysis (MA) is a statistical method that combines and analyses results from various independent studies that address the same research question, to conclude a summary estimation with more reliable comparison among those individual studies [155]. The method is widely applied in different disciplines and sciences such as education

[156], psychology [157, 104], epidemiology [158], criminology and criminal justice [157]. Additionally, it is frequently used as an overall assessment of the diagnostic test accuracy across populations and environments [104].

The primary purpose of the MA in the diagnostic test accuracy (DTA) studies is to pool sensitivity (Se) and specificity (Sp). The calculation of the pooled Se and Sp involves combining, or aggregating data from different groups or studies to derive summary values representing the overall sensitivity and specificity. This pooling combines the sensitivity (Se) and specificity (Sp) from the independent studies in order to estimate their combined or pooled sensitivity and specificity [151]. The pooling is more than just calculating the average of Se and Sp values, it requires statistical models to integrate these values, accounting for each state contribution based on factors like sample size and heterogeneity [155].

A number of statistical models are utilised in the MA of the DTA, including separate or joint models, along with fixed- or random-effects models [155]. The separate pooling models (univariate models) are rare and not recommended because they fail to consider the correlation between sensitivity and specificity [159], where these measures often have a negative correlation within studies [160]. However, these models are frequently used for analysing either Se or Sp, but not both simultaneously. Moreover, they can be applied when estimating all parameters of the joint model is challenging [159].

On the other hand, joint pooling of the sensitivity and specificity can be achieved through bivariate models or the hierarchical summary receiver operating characteristic (HSROC) model, which involves the simultaneous estimation of sensitivity and specificity as two correlated outcomes. Both models structures conduct data analysis by considering two levels; level one models a within-study variability (i.e., random sampling error) and level two models a between-study difference (i.e., heterogeneity) [155]. Moreover, the bivariate model can be employed when the independent studies or groups have minimal variation in these measures (Se and Sp), and usually occurs when included studies use the same diagnostic threshold or criteria and uniform treatment protocols [155]. In most situations, the bivariate random-effects model is more appropriate as it accounts for the variability between studies and allows for more generalisable findings [161].

Bivariate random-effects meta-analysis model (BRMA)

The bivariate random-effects meta-analysis model (BRMA) was proposed by Reitsma et al. in 2005 [160]. Their model is a general linear mixed model and includes two levels representing within (level one) and between (level two) study variability. Level one of the model makes an assumption of an approximate normal distribution for the observed logit sensitivity and logit specificity. It is based on the logit transformations of sensitivity and specificity, because the logit is on an unbounded continuous scale and can assume to follow a normal distribution [162]. In contrast, level two of the model enables a joint distribution through a bivariate normal distribution of the logit sensitivities and logit specificities by integrating two correlated normal distributions [160].

The model proposed by Chu et al. (2006) [160] contributes a significant modification in the Reitsma et al. framework. The Chu et al. model introduces a distinct approach at level one. It addresses the within-study sampling variability (level one) by using an exact binormal distribution for the sensitivity and specificity in each study, while the random-effect (level two) remains as in the Reitsma et al. model. Therefore, the Chu et al. model is a bivariate generalized linear random-effects model. This approach yields unbiased estimation, unlike the general linear random-effect model by Reitsma et al. which may provide biased estimates. This is an improvement because the Chu et al. model eliminates the need for ad hoc corrections and provides reliable estimations even with the small sample sizes of both diseased and non-diseased groups. Therefore, the bivariate generalized linear random-effects model of MA has been a cornerstone framework since 2006 and remains a crucial framework in the current research. For example, in January 2024, Shi et al. [163] employed a bivariate generalized linear random-effects model for 21 studies to estimate pooled sensitivity and specificity for the diagnostic performance of machine learning models in diagnosing early gastric cancer based on endoscopic images.

4.1.2 Study objective

The main goal of this study is to extend the approach from Chapter 3 to improve the estimates of the sensitivity and specificity of the Plantix app for detecting FAW in the maize. This estimate is made within a single dataset and in the absence of a gold standard or alternative methods for comparison. Therefore, a statistical framework is introduced in this study to handle the challenges of evaluating diagnostic test

performance. Hence, by addressing this objective, the research offers a comprehensive evaluation of the Plantix app's performance in FAW detection. In addition, it can contribute to the enhancement of its reliability for farmers, agricultural and research users.

The remainder of the chapter is as follows. In section 4.2, the detailes of the method used to achieve the study's objective are presented. Section 4.3 offers a presentation of the results obtained from the study, and section 4.4 includes some discussion of the significance of the results while also acknowledging the limitations of the study.

4.2 Materials and methods

4.2.1 Method for extracting data from the mixture models

The data used in this chapter were extracted from the fitted mixture models (see Figure 3.5) using point estimates (see Table 3.3) in Chapter 3 . The extracted data was utilised to build a 2×2 diagnostic table for each Indian state, where the table was specifically constructed around the presence or absence of FAW. To achieve this, the ArcGIS Pro, a comprehensive geographic information system (GIS) software suite developed by Esri, was used to extract the Plantix app notifications within each state in India, Figure 4.1. Based on these notifications and the fitted mixture models using point estimates, the expected number of F and T cases were calculated for each state to build up a 2×2 diagnostic tables.

The process of extracting the TP and TN observations within each state was based on the number of observation $frequency_i$ at each similarity rate ($1 \leq i \leqslant 100$) and the fitted curve $\widehat{\gamma}_i$of the true normal distribution (see Eq. 3.2). The following two steps were applied to count the TP cases for FAW category and TN for Healthy category and Non-FAW category:

- Step 1, the expected number of observations belonging to the fitted curve of the T distribution $\widehat{\gamma}_i$was calculated at each similarity rate ($1 \leq i \leqslant 100$) using the formula:

$$E\left[frequency_i\right] = frequency_i \times (\widehat{\gamma}_i) \tag{4.1}$$

$$\text{where } \widehat{\gamma}_i \text{is defined in Eq 3.2}$$

- Step 2, the total numbers of T cases for each state was calculated by summing the expected numbers in first step.

- Step 3, the above two steps were applied to true fitted mixture model of FAW to extract TP, to true fitted mixture model of Healthy to extract TN and to true fitted mixture model of Non-FAW to extract TN.

However, the direct extraction of the FP and FN cases was a challenge from the fitted False (F) curves. This was because the False curves captured a mix of two different cases. For example, the FN for the Healthy category indicates that the Plantix app stated a tested maize as Healthy while the mixture model suggested to be not Healthy. Therefore, since we have three categories (FAW, Healthy, Non-FAW), the case belongs to one of the two remaining categories: either FAW or Non-FAW. The FN for the Non-FAW category indicates that the Plantix app stated a tested maize as Non-FAW while the mixture model suggested to be not Non-FAW. Thus, the case belongs to either FAW or Healthy. Moreover, the FP curve picked up Healthy and non-FAW cases. The true curves were easier to define, for example, the TP curve included only maize infested with FAW.

Therefore, to count the expected number of False cases of each category, the mixture models with point estimates and the highest similarity (top-1 similarity) and the second-highest similarity (top-2 similarity) from the Plantix app dataset were used. As discussed earlier in Chapter 3, the diagnosis with the top-1 similarity was considered to be the most accurate and reliable. The top-2 similarity diagnosis was also worthwhile to determine the expected number of each singular category under the F curves. Both the top-1 and the top-2 similarity diagnoses must have different statuses where they cannot share the same diagnostic outcome. Hence, if the top-1 similarity was under FAW category, the top-2 similarity should explore the two remaining possibilities groups.

The method was implemented individually for each category and for each state as did with True cases, emphasising that the category was identified based on the fitted mixture model for the variable top-1 similarity. The steps below were applied:

- Step 1, the category within the top-1 similarity was split into two subgroups based on the binary categories within the top-2 similarity.

  - For example, if FAW is identified under the top-1 similarity, the top-2 similarity provides two subgroups: one includes FAW and Healthy, and the other

includes FAW and Non-FAW

- Step 2, within each subgroup, the cases were merged and counted based on the top-1 similarity value to have $frequency_i$ for $(1 \leq i \leq 100)$.
  - For example the subgroup with including FAW and Healthy, at each top-1 similarity value $(1 \leq i \leq 100)$ the number of observations $frequency_i$ was determined.
- Step 3, the following formula was applied to step 2 in order to calculate the expected frequency for each i within each subgroup individually:

$$E\left[frequency_i\right] = frequency_i \times (1 - \widehat{\gamma}_i)$$

- Step 4, the resulting numbers for each subgroup from Step 3 were combined to provide the total number of False cases.

Figure 4.1: The study Area is the colored region, including South states (dark brown) and North states (light yellow).

4.2.2 Study assumptions

To estimate the combined Se and Sp of detecting FAW on infested maize crops at state-level in India, using Bayesian approach, the following assumptions were considered:

1. The diagnostic accuracy of the Plantix app and the mixture models (chapter 3) is below 100% and relatively low.
2. The study assumed that the sensitivity and specificity varied across states.
3. The study assumed that the data collected from each state were independent and identically distributed within that state.
4. The study employs a random effects model to account for the variability between states, allowing for more accurate estimation of the overall diagnostic test accuracy.

4.2.3 Criteria for selecting states

To effectively estimate the combined sensitivity and specificity, it is essential for each state to satisfy two criteria.

- The observed sensitivity for correctly detecting FAW (Se_i) and the observed specificity for correctly detecting not having FAW (Sp_i) in each state should satisfy the Yadon index (J) criterion, which states that $(Se_i + Sp_i) - 1 > 0$. According to Flor et al. (2020), the negative value of J indicates a random diagnosis so that the observed outcomes are obtained by chance rather than through a reliable diagnostic procedure [164]. Additionally, Toft et al. (2007) [127] stated that the negative value of J is not a practical value.
- The sample size should be high, otherwise the estimation is biased, and the variance may increase [165]. Moreover, Teare et al. [165] highlighted that an inadequate sample size may lead to either overestimation or underestimation of the test performance. The sample size here was selected to be greater than 30 based on the Central Limit Theorem (CLT), ensuring the sampling distribution of sensitivity (Se) and specificity (Sp) approximates a normal distribution, enabling robust and unbiased inferences. Furthermore, sample size above 30 provides sufficient data to approximate the binomial distribution (the likelihood of the Bayesian

model, see below 4.2.6) for accurate calculation of proportions. Additionally, with non-informative priors, a larger sample size enhances the likelihood's influence on the posterior distribution, making the estimates more data-driven and compatible with the bivariate normal prior for logit-transformed Se and Sp.

4.2.4 Fixed-effect (FE) vs. random-effects (RE) models in meta-analysis

In the meta-analysis are two well-known models are used; the fixed-effect model (FE) and the random-effects model (RE). The two models differ in their unique assumptions based on the true effect size across studies. The effect sizes are quantitative values that estimate the extent of the difference or association among interested variables. In the context of the MA of DTA, the effect size is the differences between the true values of sensitivity (or specificity) and their observed values in individual studies.

Fixed-effect model

The fixed-effect (FE) model relies on the assumption that there is a unique true effect size that is common or fixed across all studies encompassed in the analysis. Therefore, the observed diference between individual studies is only due to sampling error (random chance in selecting units). Hence, this model is also known as a common-effect model [166]. However, when studies are conducted in diverse environments with different government regulations, languages, and cultural contexts, the assumption of homogeneity may be violated. In such cases, applying a fixed-effects model may lead to an overly optimistic estimate of the treatment effect, as it does not adequately account for the inherent differences among studies. This limitation can result in misleading conclusions regarding the generalizability of the findings.

Random-effect model

With the random-effects (RE) model, on the other hand, the true effect sizes can differ between studies, due to sampling error and true variability in population parameters. For example, at the state level, the effect size can change due to the differences in farmers skill in taking images and the quality of the smartphones in terms of internet connection or camera resolution. Moreover, the effect size can differ due to environmental reasons, for example, the maize growth stage and severity of the infestation. The RE model

estimates both the average effect size across studies and the variance in true effect sizes between them [166].

4.2.5 Bivariate generalized linear mixed effects model of meta-analysis

Suppose that there were n independent studies (states) in test accuracy of the Plantix app, labeled from $i = 1$ to n. Each single state had maize crops infested with FAW ($n_{i,\,FAW}$) and non infested with FAW ($n_{i,\,\overline{FAW}}$), where $n_{i,\,FAW}$ included FAW cases that were correctly identified (TP_i) and instances that were missed to recognise FAW (FN_i). While, $n_{i,\,\overline{FAW}}$ addressed instances Healthy or non-FAW cases. It consisted of accurately recognised Healthy or Non-FAW cases (TN_i) and instances that incorrectly identified as FAW (FP_i). Therefore, the observed sensitivity in each study was $Se_i = \frac{TP_i}{TP_i+FN_i} = \frac{TP_i}{n_{i,\,FAW}}$, and the observed specificity was $Sp_i = \frac{TN_i}{TN_i+FP_i} = \frac{TN_i}{n_{i,\,\overline{FAW}}}$.

To build the bivariate model, the generalized linear mixed model with two levels was employed, as proposed by Reitsma et al. and adapted by Chu et al. [160]. At level one (within-study variability), the number of true positive (TP_i) and the number of true negative (TN_i) respectively, were assumed to follow exact binomial distribution at each study, as suggested by Chu et al. [160]. That is:

$$\begin{aligned} TP_i &\sim Bin(n_{i,\,FAW}, Se_i). \\ TN_i &\sim Bin(n_{i,\,\overline{FAW}}, Sp_i). \end{aligned} \tag{4.2}$$

The binomial distribution is a better assumption than the normal distribution that was proposed by Reitsma et al. for the following reasons: (i) the sample size can be small, (ii) Se and Sp are limited to the interval $[0, 1]$, and (iii) an ad hoc continuity correction (adding 0.5) is not required when any of the counts in the 2×2 diagnostic table are zero in a state [160]. (iv) the binomial distribution is also appropriate for describing binary outcomes such as presence or absence of FAW. While the binomial distribution is suitable for small sample sizes, and small sensitivity and specificity values, it is important to acknowledge the above criteria.

To account for between-study heterogeneity (level two), each Se_i and Sp_i converts to the logit scale. Then, the true logit-transformed sensitivities, $logit(Se_i) = log(\frac{Se_i}{1-Se_i})$ and specificities, $logit(Sp_i) = log(\frac{Sp_i}{1-Sp_i})$ are assumed to follow a bivariate normal

distribution as proposed by Reitsma et al. [160], and this distribution captures between-study correlation. It takes the form,

$$logit \begin{bmatrix} Se_i \\ Sp_i \end{bmatrix} \sim MVN \left[\boldsymbol{\beta} = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}, \quad \Omega \right], \Omega = \begin{bmatrix} \tau_1^2 & \tau_1\tau_2\rho \\ \tau_1\tau_2\rho & \tau_2^2 \end{bmatrix}, \tag{4.3}$$

where β_1 is the mean of the logit sensitivities, and β_2 represents the pooled specificity on the logit scale. Hence, the pooled sensitivity can be calculated using the inverse of the mean logit transformation sensitivities and can be expressed as pooled $Se = logit^{-1}(\beta_1) = \frac{e^{\beta_1}}{1+e^{\beta_1}}$. The pooled specificity can be calculated using the inverse of the mean logit transformation specificities and can be expressed as pooled $Sp = logit^{-1}(\beta_2) = \frac{e^{\beta_2}}{1+e^{\beta_2}}$. A variance-covariance Matrix is Ω. The diagonal parts of the variance-covariance matrix $\left(\tau_1^2, \tau_2^2\right)$, the variance quantifies the spread or variability of the logit sensitivities τ_1^2 and the logit specificites τ_2^2. On the other hand, the off-diagonal term $(\tau_1\tau_2\rho)$ describes pairwise covariances between logit sensitivities and logit specificities, where ρ represent a correlation coefficient between both terms. The covariance term reflects the power and trend of the linear relationship between each of them and it is limited to the interval $[-1, 1]$, where zero covariance indicates no linear relationship between them. Therefore, there are five different parameters that need to be estimated. Hence, at least five independent states should be used to identify these parameters [161]. Estimation of the above model can be achieved in both the classical and Bayesian frameworks [167]. This work focuses on Bayesian approach with the MCMC algorithm.

4.2.6 MCMC for the joint bivariate normal distribution of logit Se and Sp

The Bayesian approach combines prior and likelihood to obtain the posterior distribution, and MCMC algorithm was applied to estimate all required parameters. The likelihood component included the two independent binomial distributions, as provided in equation (4.2), where the TP, FP, TN and FN were estimated using equations 4.1 and 4.2.1. After the likelihood had been determined, two vague prior distributions were outlined for the mean vector $\boldsymbol{\beta}$ and the covariance matrix Ω parameters. For $\boldsymbol{\beta}$, a multivariate normal distribution (MVN) was employed as a prior:

$$\boldsymbol{\beta} = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} \sim MVN \left[\boldsymbol{\mu} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \Sigma = \begin{bmatrix} 0.1 & 0 \\ 0 & 0.1 \end{bmatrix} \right].$$

In this instance, it was a weak informative prior, because there was no prior knowledge of $\boldsymbol{\beta}$, and the mean vector μ, and the precision matrix Σ, were vague.

Here μ was represented by zeros to be unbiased towards certain values. Further, the Σ is known as the precision matrix, and it is an inverse of variance-covariance matrix. In other words, the precision matrix represents the inverse of variance (precision) [167]. Consequently, to reflect vague or non-informative prior of Σ, comparatively large values for the variances were selected for the diagonal to allow for a large variety of potential values. Therefore, value of 10 was chosen here as variances, equivalent to 0.1 ($\frac{1}{10}$) in the precision matrix. Further, small values in the off-diagonal elements indicate no correlation between the variables.

For the covariance matrix the inverse-Wishart distribution is often considered as prior distribution [167, 168]. The inverse-Wishart distribution is defined with two parameters, namely the scale matrix (S) and the degrees of freedom (df) [167, 168]. The scale matrix is frequently initialised to be an identity matrix [167, 168], containing ones along the diagonal and zeros in the off-diagonal. As a result, the identity matrix is a practical selection [168] for a prior distribution when there is no prior knowledge. The df represents the amount of information available for estimating the parameters in the model. Since there is no prior knowledge, df was chosen as the smallest possible value, which is 2. The lowest df increases the uncertainty, leading the model to be more flexible in its estimates and better capturing of the data variability [167].

$$\boldsymbol{\Omega}^{-1} \sim \text{ Wish } \left(S = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, df = 2 \right).$$

The model was fitted in WinBUGS14 software [169] and the R2WinBUGS package [170] to estimate the joint and marginal posterior distributions of the model parameters. The Markov Chain Monte Carlo (MCMC) algorithm was run for 10,000 iterations. In addition, a warm-up period of the initial 3000 iterations were discarded to enable the samples to reach the stationary distribution. Three chains were applied to guarantee convergence. Subsequently, the trace plots, density plots, effective sample size (ESS), Gelman-Rubin statistics ($\hat{R}$) and autocorrelation plots were checked for convergence of the posteriors and marginal posterior distributions.

After that, posterior predictive check was applied, examining whether the model produces samples that can represent the observed data. This procedure facilitated the assessment of the suitability of the model to fit and explain the data. Therefore, after excluding the initial 3000 samples, the histograms of the posterior distributions of sensitivity and specificity iterations, with dashed line of observed sensitivity and specificity from the data for each states were drawn. If the observed value alignd closely to the peak of the histogram (expected of the sensitivity or specificity), it was considered to be a good indicator that the model reflected the data well and yielded effective findings.

4.3 Results

4.3.1 State-level data

The method was used to build a 2×2 diagnostic test table for each state. The Table 4.1 summarises the 2×2 diagnostic test table at state level for the detection of the FAW when the Healthy category as negative, and the Table 4.2 represents the results with Non-FAW group as the negative. The data in both Tables is sorted by the Yadon Index value (J), starting with the highest value.

Table 4.1: Summary of state-level data for FAW as positive and Healthy as negative. TP means number of true positives, FP indicates number of false positives, TN reflects number of true negatives, and FN means number of false-negatives. P is the total number of positives, mathematically (P=TP+FP) and the Total indicates the total number of tested cases, mathematically (Total=TP+FP+TN+FN). Se, Sp and J indicates sensitivity, specificity and Yadon Index, respectively.

State	TP	FP	TN	FN	P	Total	Se	Sp	J
Mizoram	2	0	3	0	2	5	1	1	1
Nagaland	5	0	5	1	5	11	0.83	1	0.83
Telangana	1381	256	1057	685	1637	3379	0.67	0.81	0.48
Maharashtra	4004	719	2631	1901	4723	9255	0.68	0.79	0.47
Manipur	9	2	6	4	11	21	0.69	0.75	0.44
Karnataka	2209	499	1846	1277	2708	5831	0.63	0.79	0.42
Andhra Pradesh	1340	330	949	648	1670	3267	0.67	0.74	0.41
Bihar	364	108	896	375	472	1743	0.49	0.89	0.38
Tamil Nadu	575	152	316	303	727	1346	0.65	0.68	0.33
Gujarat	356	134	381	291	490	1162	0.55	0.74	0.29
Madhya Pradesh	932	437	1607	992	1369	3968	0.48	0.79	0.27
West Bengal	186	83	230	189	269	688	0.5	0.73	0.23
Jharkhand	82	39	205	127	121	453	0.39	0.84	0.23
Odisha	48	21	47	41	69	157	0.54	0.69	0.23
Assam	27	18	57	32	45	134	0.46	0.76	0.22
Chhattisgarh	164	92	264	198	256	718	0.45	0.74	0.19
Uttarakhand	24	19	46	44	43	133	0.35	0.71	0.06
Rajasthan	256	315	1090	702	571	2363	0.27	0.78	0.05
Uttar Pradesh	409	431	1557	1087	840	3484	0.27	0.78	0.05
Punjab	113	128	321	238	241	800	0.32	0.71	0.03
Arunachal Pradesh	1	2	7	4	3	14	0.2	0.78	-0.02
Himachal Pradesh	11	23	112	65	34	211	0.14	0.83	-0.03
Haryana	123	159	263	245	282	790	0.33	0.62	-0.05
Kerala	6	13	30	24	19	73	0.2	0.7	-0.1
Goa	2	3	4	4	5	13	0.29	0.57	-0.14
Tripura	2	3	3	7	5	15	0.22	0.5	-0.28

It is clear from Table 4.1 that there are some individual states where the Yadon index, (J) value is negative (light red), or the sample size is less than 30 (light yellow). Arunachal Pradesh, Himachal Pradesh, Haryana, Kerala, Tripura and Goa have negative values, whereas Mizoram, Nagaland, and Manipur have small sample size, of less than 30, which may lead to misrepresentation of the true performance of the Plantix app test. For example, a sensitivity of 1 for the Mizoram state could be due to the limited sample size of 5, and it is a good indication that there may be an overestimation. Some states also have both a small sample size and a negative Yadon index, such as Arunachal Pradesh, Kerala and Tripura. In addition, in the rest of the 17 states, it can

be noted that there is a variability in the observed sensitivity (Se) and the observed specificity (Sp) among different states, despite the diagnostic test and the classification algorithm being exactly the same. The Table 4.1 shows the sensitivity ranging from 0.19 to 0.68. Likewise for specificity ranges were from 0.68 to 0.89.

Table 4.2: Summary of state participant data where FAW as positive and Non-FAW as negative. TP means number of true positives, FP indicates number of false positives, TN reflects number of true negatives, and FN means number of false-negatives. P is the total number of positives, mathematically (P=TP+FP) and the Total indicates the total number of tested cases, mathematically (Total=TP+FP+TN+FN). Se, Sp and J indicates sensitivity, specificity and Yadon Index, respectively.

State	TP	FP	TN	FN	P	Total	Se	Sp	J
Nagaland	5	2	2	3	7	12	0.62	0.5	0.12
Assam	27	58	51	31	85	167	0.47	0.47	-0.06
Goa	2	8	11	4	10	23	0.33	0.58	-0.09
Mizoram	2	4	8	7	6	21	0.22	0.67	-0.11
Madhya Pradesh	932	1861	979	875	2793	4647	0.52	0.34	-0.14
Odisha	48	112	66	55	160	281	0.47	0.37	-0.16
Maharashtra	4004	6109	1873	2649	10113	14635	0.6	0.23	-0.17
Uttarakhand	24	69	39	31	93	163	0.44	0.36	-0.2
Telangana	1381	2651	1056	1263	4032	6351	0.52	0.28	-0.2
Karnataka	2209	4244	1391	1988	6453	9832	0.53	0.25	-0.22
Gujarat	356	694	214	305	1050	1569	0.54	0.24	-0.22
Andhra Pradesh	1340	2632	892	1216	3972	6080	0.52	0.25	-0.23
Tamil Nadu	575	1156	300	443	1731	2474	0.56	0.21	-0.23
Manipur	9	12	5	10	21	36	0.47	0.29	-0.24
Chhattisgarh	164	376	182	216	540	938	0.43	0.33	-0.24
Jammu	9	42	37	25	51	113	0.26	0.47	-0.27
Rajasthan	256	863	402	363	1119	1884	0.41	0.32	-0.27
Himachal Pradesh	11	73	93	59	84	236	0.16	0.56	-0.28
Jharkhand	82	261	102	110	343	555	0.43	0.28	-0.29
Haryana	123	346	107	152	469	728	0.45	0.24	-0.31
Bihar	364	980	533	728	1344	2605	0.33	0.35	-0.32
West Bengal	186	479	215	312	665	1192	0.37	0.31	-0.32
Kerala	6	51	51	30	57	138	0.17	0.5	-0.33
Tripura	2	6	3	4	8	15	0.33	0.33	-0.34
Uttar Pradesh	409	1514	691	759	1923	3373	0.35	0.31	-0.34
Punjab	113	410	170	236	523	929	0.32	0.29	-0.39
Arunachal Pradesh	1	8	3	4	9	16	0.2	0.27	-0.53

On the other hand, all states in Table 4.2 have a negative J except Nagaland, which has small sample size of 12. Subsequently, the detection of FAW when Non-FAW as negative was excluded from further analysis due to its failure to satisfy the basic assumptions. The misdiagnosis of maize infested with FAW, when they were Non-FAW

as negatives, can be related to the resemblance between damage symptoms or physical characteristics among FAW and Non-FAW. The most significant misdiagnosis within the Non-FAW class occured with violet stem borer and spotted stemborer (see Figure 4.2).

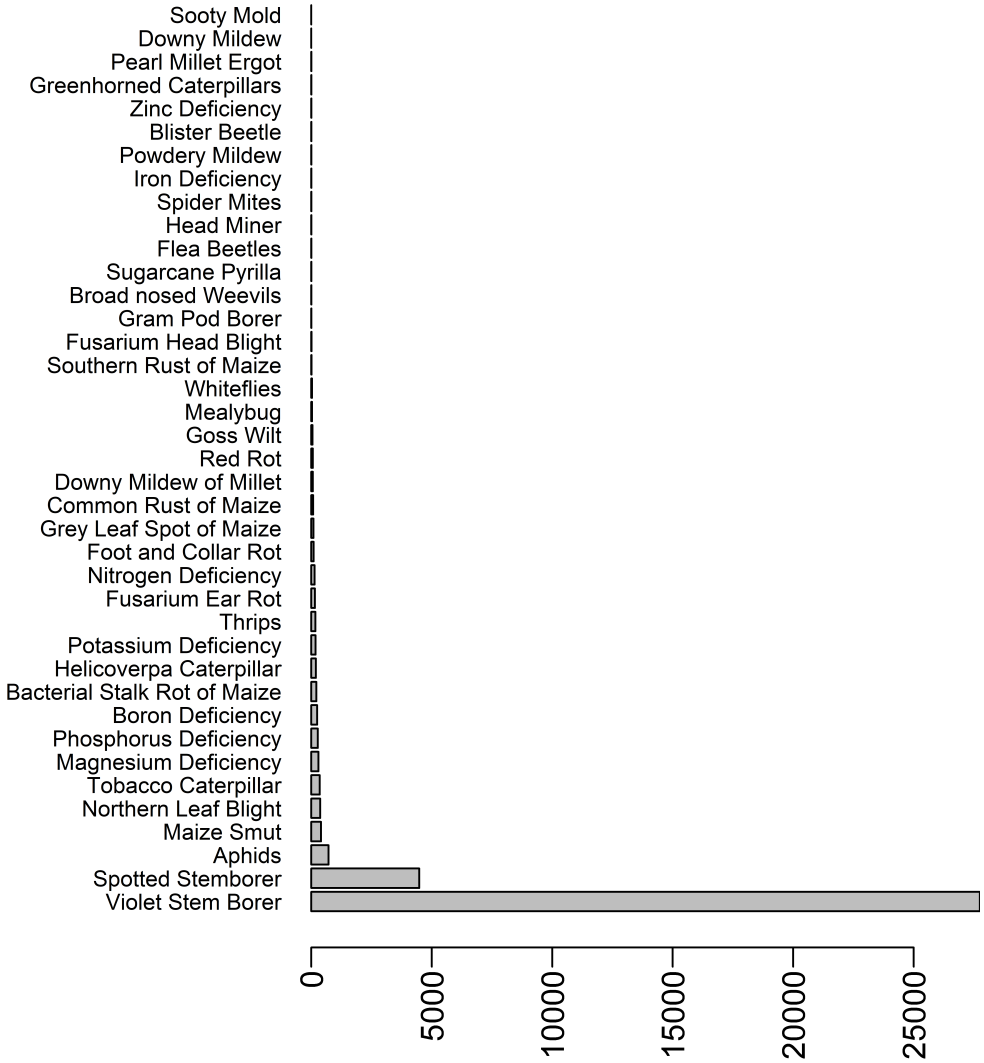


Figure 4.2: Frequency of false positive observations of Non-FAW category under the top-2 similarity.

4.3.2 MCMC for the joint model of the logit sensitivity and specificity

It is clear from Figure (4.3) that there is a variation between the sensitivity and the complement of the specificity (1-Sp) based on 17 states in Table 4.1. It is also noticeable that all states lie above the line $y = x$, indicating the Plantix app was more effective at correctly identifying true positives than incorrectly identifying false positives. In

addition, a weak positive linear correlation (Pearson Correlation=0.07) between these two variables (Se and $(1 - Sp)$) can be observed. Hence, it can be expected that the sensitivity and specificity can have a negative linear relationship. Moreover, the scatter plot follows an ellipse, supporting a possible bivariate distribution, for instance, the bivariate normal distribution. The bivariate model enables the analysis of their joint distribution. This joint bivariate normal distribution serves as the prior for sensitivity and specificity in the Bayesian framework.

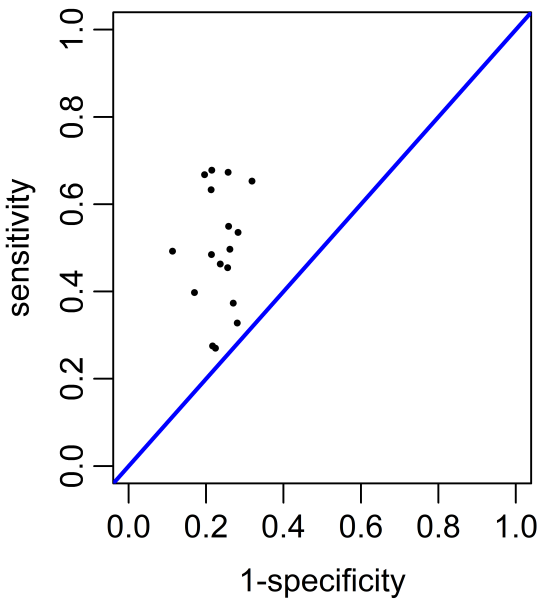


Figure 4.3: Scatter plot of the sensitivity (Se) and the complement of the specificity ($1 - Sp$) from Table 4.1. The blue line represents $y = x$, reflecting equal values of sensitivity and (1−specificity). Plotting Se against $1 - Sp$ aligns with the ROC curve.

The bivariate random-effects meta-analysis model was applied after converting sensitivity and specificity into the logit scale. The results presented in detail by including three main elements: (i) evaluating convergence, (ii) estimating parameters with their credible intervals of uncertainty, and (iii) conducting posterior predictive checks.

the convergence of three MCMC chains was assessed to ensure the posterior estimate reliability. The trace plots, density plots, ACF plots, Gelman-Rubin statistics, and effective sample sizes were used as diagnostic tools. These tools were applied after removing a 3000 burn-in period of 10,000 iterations. The trace plots of the three chains for means, variance, covariance marginal posterior distributions and Se and Sp showed convergence (Figure 4.4, left column). This is because the triple-chain trace plots for each parameter show that they overlap and exhibit similar behaviour. Therefore, they

have almost the same point estimated values as well as credible intervals. Moreover, every single chain reflects stationarity and well mixing, because there is no trend and fluctuates around a consistent value. Thus, the three MCMC chains for each parameter are expected to converge to the same distribution. Additionally, the density plots for each interested parameter were inspected, as shown in Figure 4.4, right column. Each parameter with three independent chains reflects univariate distributions and overlapping curves. Therefore, the plots indicates convergence to a common distribution.

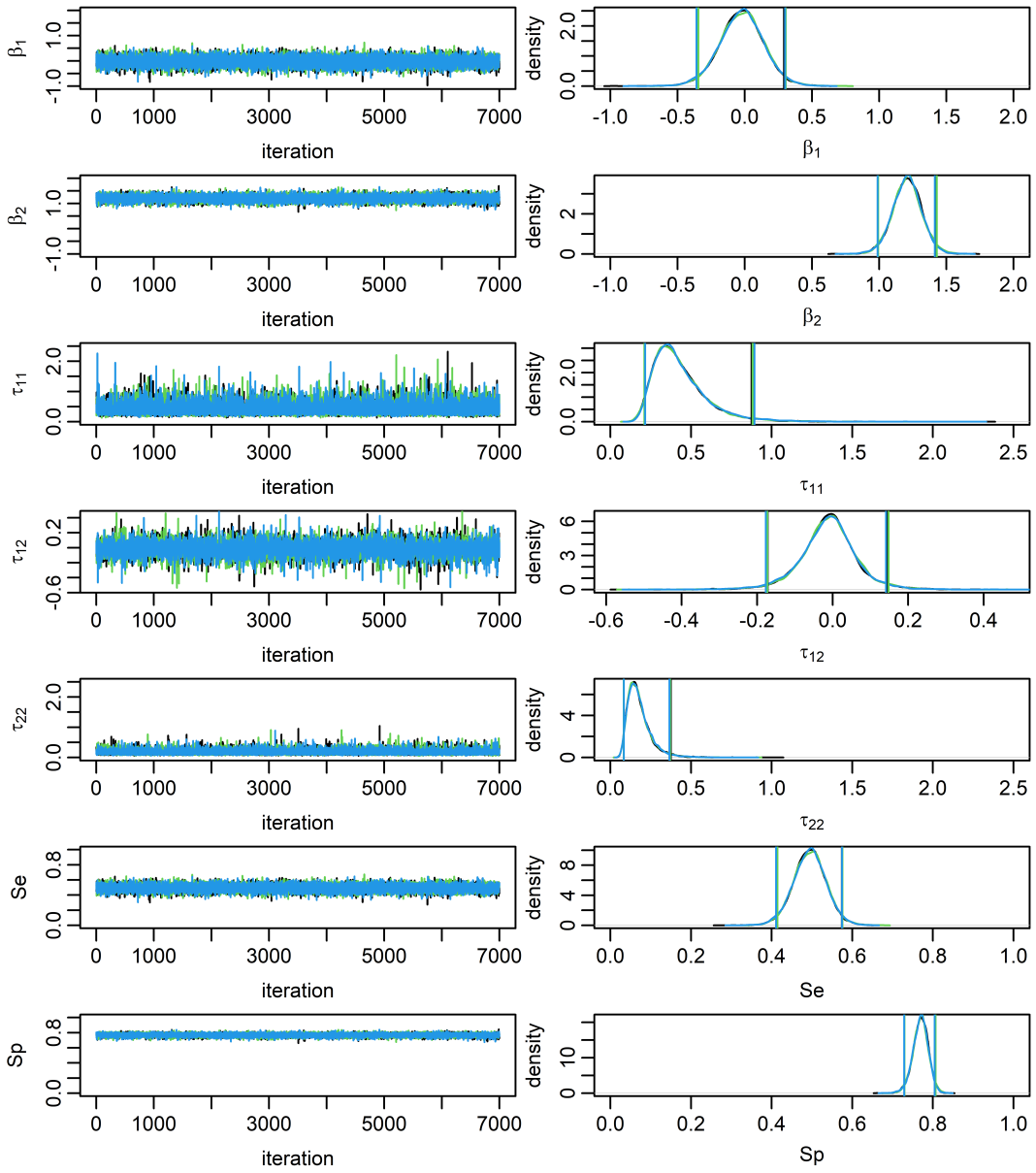


Figure 4.4: Trace (left column) and density with vertical lines to represent 95% credible intervals (right column) plots for convergence assessment of the three MCMC chains for model parameters (β_1,β_2, τ_{11},τ_{12}, τ_{22} ,Se, Sp).

As indicated in Figure 4.5 the autocorrelation function (ACF) plots which represent the correlation within MCMC samples of each parameter, confirm convergence. This is because the ACF plots display rapid decay of autocorrelations and fluctuation near zero. Therefore, the plots indicate that successive samples (iterations) are independent and hence demonstrate good mixing. As a result, the MCMC chains are sufficient to define the posterior distributions.

While ACF plots provide a graphical assessment of autocorrelation, effective sample size (ESS) summarises the overall efficiency through a quantitative value. Therefore, the ESS was calculated, and the values for each model parameter were found to be 21000, which is equivalent to the total number of iterations across all three chains after excluding the burn-in period (7000×3). As a result, the ratio of the ESS to the total number of iterations ($\frac{\text{ESS}}{21000}$) for each model parameter was 100 percent, indicating minimal correlation between successive samples (see subsection 2.5.1 in Chapter 2). Hence, all MCMC iterations generated independent samples and proved to be valuable in reliably estimating the distribution.

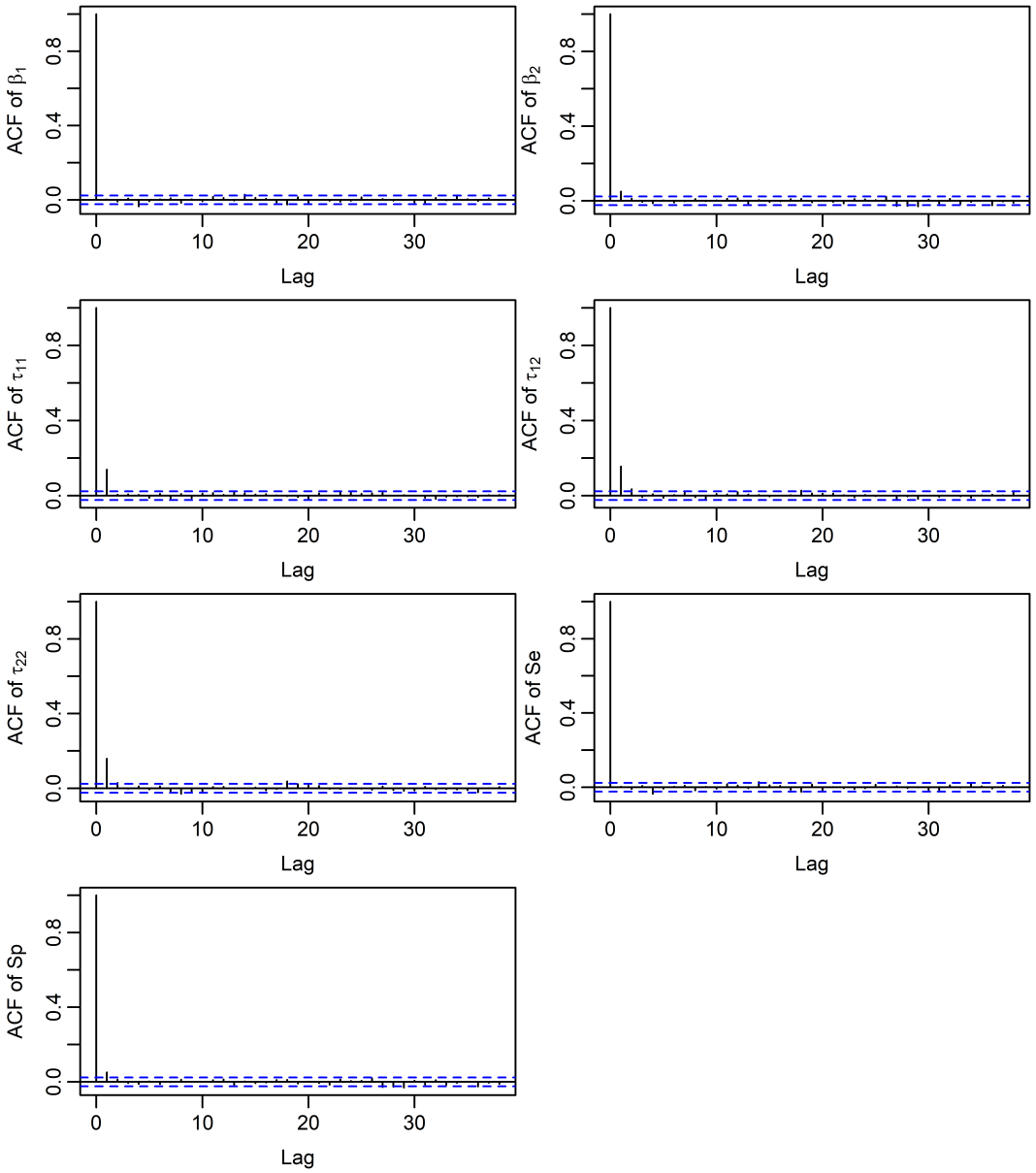


Figure 4.5: Autocorrelation function (ACF) plot for the convergence assessment of MCMC

In addition, the Gelman-Rubin statistic (R-hat) value was applied to all three chains of each marginal posterior distribution. The R-hat value was equal to one for each parameter. As a result, the within-chain variance was equal to the between-chain variance. Therefore, the chains reached convergence and provided reliable and consistent posterior estimates. To sum up, based on this analysis, the MCMC chains had reached convergence. Thus, they provided reliable posterior samples.

The logit transformation model and estimated values for the parameters are

$$logit \begin{bmatrix} Se_i \\ Sp_i \end{bmatrix} \sim MVN \left[\boldsymbol{\beta} = \begin{pmatrix} -0.023 \\ 1.208 \end{pmatrix}, \begin{bmatrix} 0.434 & -0.012 \\ -0.012 & 0.179 \end{bmatrix} \right]$$

where the mean logit-sensitivities (95% credible interval) and mean logit-specificities were -0.023 (-0.352, 0.300) and 1.208 (0.992, 1.423), respectively. Moreover, the unexplained between-states heterogeneity in logit(Se) and logit(Sp) was measured as 0.434 (0.212, 0.861) and 0.179 (0.084,0.365), respectively. The posterior covariance between logit(Se) and logit(Sp) was -0.012 (-0.179,0.139). Hence, the $\rho \approx -0.043$ indicates that the relationship between sensitivity and specificity measures across states was weak and negative (see also Figure 4.6). This suggests that the effectiveness of the Plantix app in detecting the presence of FAW (sensitivity) does not significantly impact its ability to correctly identify the absence of FAW. Additionally, the pooled of Se and Sp can be obtained from the logit scale by applying the inverse logit transformation, which can be expressed as $\frac{exp(\beta_j)}{(1+exp(\beta_j)}$ where $j = \{1, 2\}$. Therefore, the estimated mean and uncertainty for the pooled sensitivity and specificity are shown in Table (4.3):

Table 4.3: Summary of estimated Parameters and uncertainty (95% credible intervals, CI) for pooled sensitivity (Se) and specificity (Sp).

Parameter	Mean= $\frac{exp(\beta_j)}{(1+exp(\beta_j)}, j = \{1, 2\}$	95% CI
pooled Se	0.494	(0.413,0.574)
pooled Sp	0.769	(0.730, 0.806)

The Table 4.3 showed that the Se was lower than the Sp. Moreover, the uncertainty in the credible intervals was narrow. This was an indication that the estimation of the parameters was precise and more reliable. Therefore, the confidence in the results could increase and enhance the robustness of the inference drawn from the analysis.

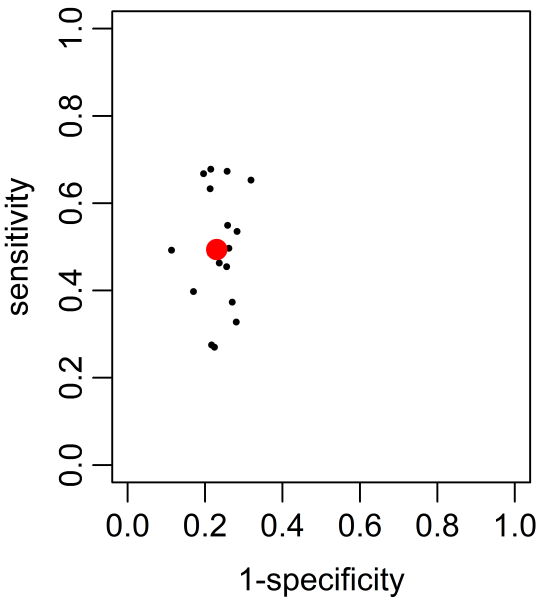


Figure 4.6: Scatter plot of the sensitivity and the complement of the specificity $(1 - Sp)$ (black points) with estimated pooled Sensitivity and the complement of the specificity $(1 - Sp)$ (red point).

Further, Figure 4.7 shows the forest plots for the Se and Sp for each state, and the pooled values at the bottom of the plots. Sensitivity values show greater variability across states comparing with specificity values. In addition, states in the top of the Se forest plot show the highest sensitivity, where these states have the highest sample size as well as the known states with highest maize producing. On the other hand, the bottom states reflect the opposit where also these states have the lowest dample size, and less maize producing. The values of specificity appear to be relatively consistent across most states, with values clustering near the upper range of the scale (around 0.8 to 1.0). High specificity suggests that the system or process being analysed is generally effective at correctly identifying true negatives across all states. The variability in sensitivity indicates that there may be challenges in accurately identifying true positives in certain states, potentially necessitating targeted interventions or adjustments to improve performance in those states.

Additionally, posterior predictive checks was applied, examining if the logit model produced samples that represented the observed Se and Sp from the data (see Se and Sp columns in Table 4.1). This application helped in evaluating the goodness of the model to fit and explain the data. Therefore, the histograms of the posterior distributions of sensitivity and specificity samples were compared with the observed sensitivity (see

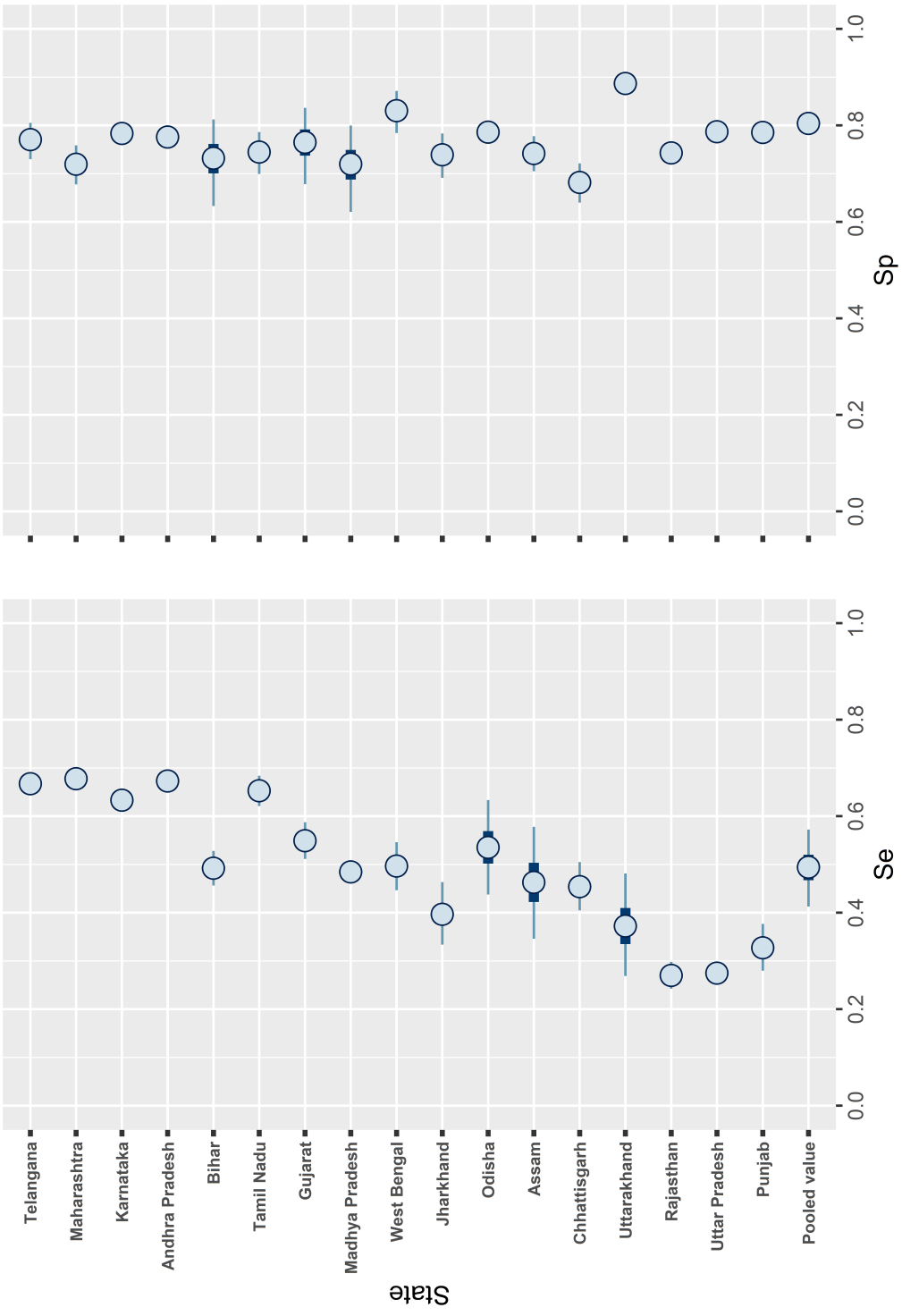


Figure 4.7: Forest plots of the mean posterior estimates for sensitivity (Se) and specificity (Sp) along with their credible intervals (CrI). The bold line represents the 50% CrI, while the thin line indicates the 95% CrI.

Figure 4.8) and specificity (see Figure 4.9) for each of the states were drawn. It can be seen that in all seventeen histograms the average sensitivity and specificity of the model aligned closely to the observed value. Thus, it can be concluded that the model was appropriate and fitted the data well.

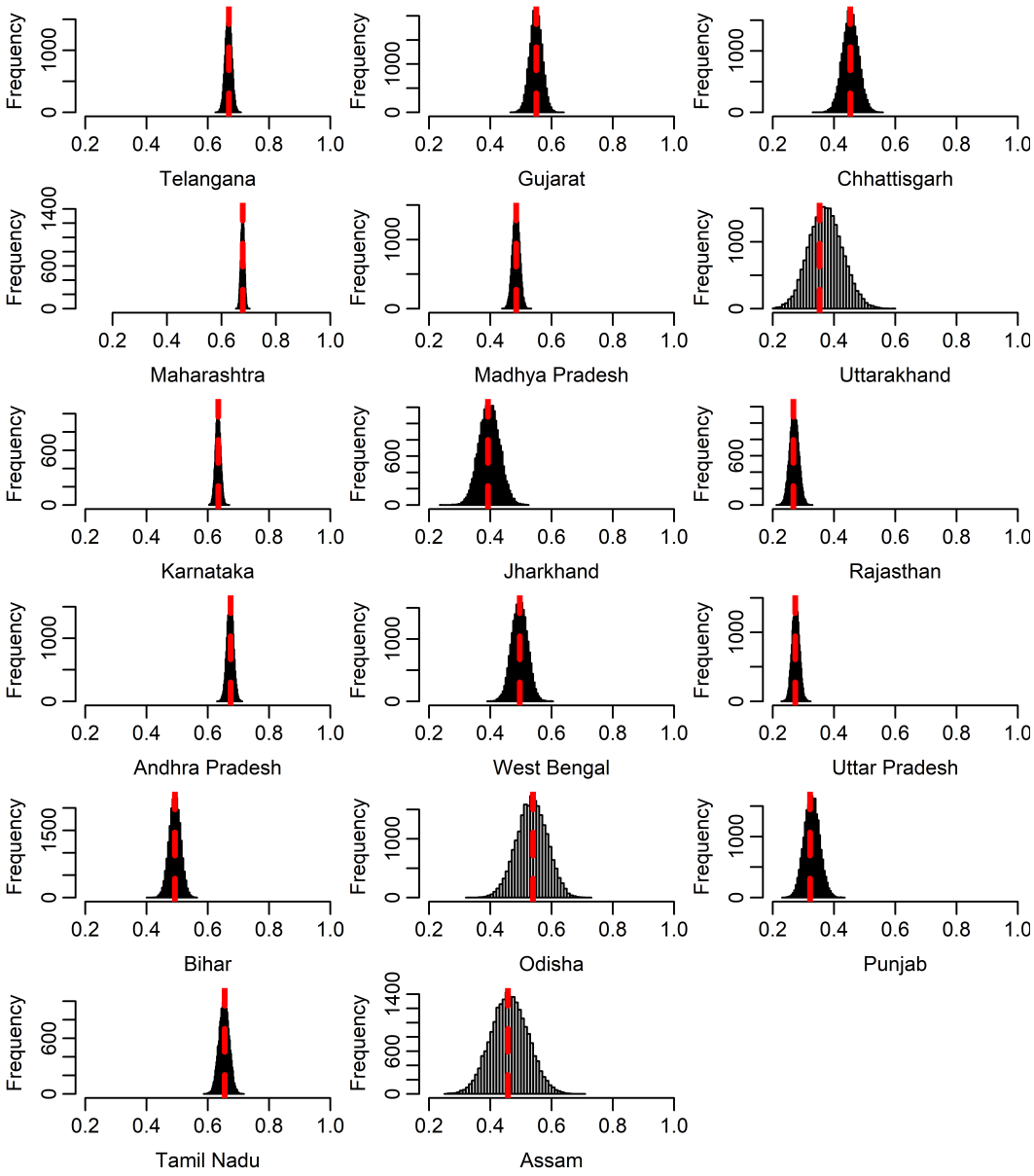


Figure 4.8: The posterior distributions of sensitivity in each Indian-state with the dashed red line representing the observed sensitivity for each state.

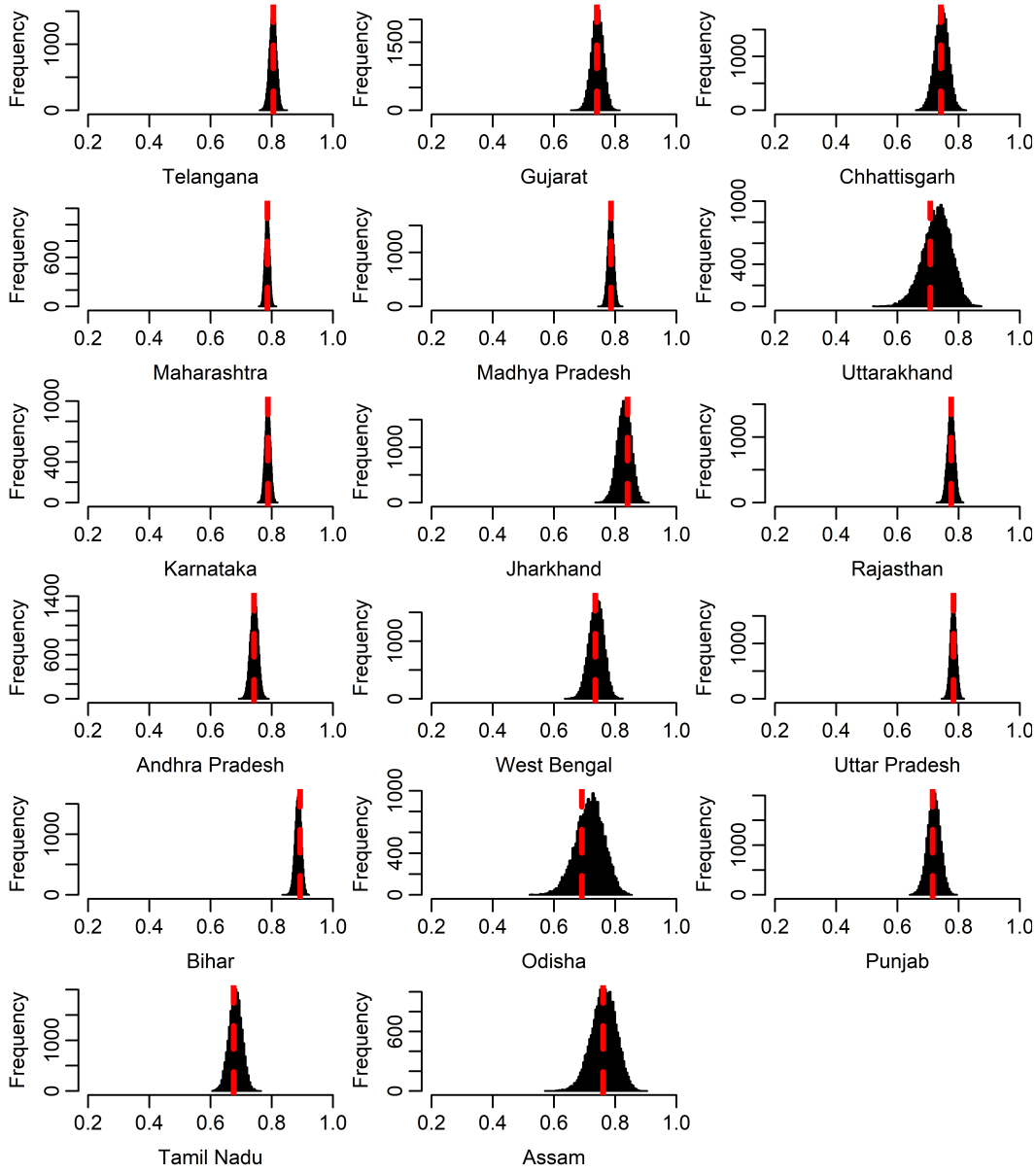


Figure 4.9: The posterior distributions of specificity in each Indian-state with the dashed red line representing the observed specificity for every state.

4.4 Discussion

The present chapter was primarily aimed to statistically estimating the pooled sensitivity and specificity of the Plantix app, the AI diagnostic test for identifying FAW in Indian maize crop. This study was based on data collected by citizen scientists via the Plantix app, where a gold standard was unavailable. Moreover, there was no evidence in literature about the app accuracy, poseing a unique challenge in evaluating the sensitivity and specificity of the Plantix app, and similar approaches in machine learning

and citizen science.

To achieve this, firstly, a 2x2 diagnostic table was constructed for each Indian state from the distinct mixture models of positive and negative notifications of the Fall Armyworm (FAW). Secondly, the meta-analysis using the bivariate generalized linear mixed effects model, with the Markov Chain Monte Carlo (MCMC) algorithm was employed. The model is based on the bivariate RE model by Reitsma et al. (2005), and the adapted bivariate generalized linear mixed-effects model. Thirdly, the model was only applied for detecting FAW when Healthy was used as negative reading, to represent the performance of the diagnostic accuracy. This is because the data for Indian states for detecting FAW when Non-FAW as negative failed to satisfy the criteria of the Yadon index and/or sample size.

The 2x2 diagnostic table in Table 4.1 reflects that nine states with the highest number of tested maize crop locate in the major maize growing states, which are Karnataka (14.8 percent), Maharashtra (10.9 percent), Madhya Pradesh (10.8 percen), Rajasthan (10.6 percent), Andhra Pradesh (10.4 percent), Uttar Pradesh (8.3 percent), Bihar (7.9 percent), Gujarat (5.0 percent) and Tamil Nadu (3.6 percent) [24]. Hence, these states with larger sample sizes contribute more to the overall pooled estimate. The results of this study demonstrate the accuracy of the Plantix app diagnostic test. For the 17 states combined, the pooled sensitivity was 0.494 (95% credible interval: 0.413-0.574). These estimates suggest that the app test has a sensitivity in detecting the FAW when Healthy as negatives, with approximately 50 percent of true positive cases being correctly identified by the test. For the specificity, the Plantix app had the high specificity of 0.769 (0.730, 0.806), indicating its strong ability to identify Healthy maize crops, while around 20 percent of Healthy crops were misdiagnosied as infested with FAW (false positives).

In addition, the correlation between Se and Sp on the logit scale was negative and weak. This weak correlation indicates that the Plantix app can successfully achieve high value in both sensitivity and specificity at the same time, with limited drawbacks. This finding holds significant value for the Plantix developers where improving the app Se does not significantly impact the app Sp. Further, the high false negative and low Se are not solely because of the ability to distinguish between FAW and Healthy crops, while this also suggests the impact of other factors such as image quality and infestation variability. Furthermore, the bivariate model is still appropriate, even with

a weak correlation. it can effectively illustrate the trend of the relationship between these two metrics and takes into account any possible interactions. Hence, it leads to reliable estimates of the diagnostic performance of the Plantix app.

The misdiagnosis of maize infested with FAW, when they were Non-FAW as negatives, can be related to the resemblance between damage symptoms or physical characteristics among FAW and Non-FAW. The most significant misdiagnosis within the Non-FAW class occured with violet stem borer and spotted stemborer (see Figure 4.2). To some extent, the FAW and these two pests have similar physical characteristics in their eggs colour as well as their shape and mass. Spotted stemborer lays eggs in masses and on the lower surface of leaves. Yet, FAW lays eggs in masses but on the upper or under side of leaves. However, violet stem borer lays eggs in two to four rows inside the leaf sheath [3]. Therefore, the physical characteristics may be similar, while the area to find those characteristics in the crop may be different. Additionally, all the three insects make holes in maize leaves. On the other hand, both FAW and spotted stemborer create a papery (transparent) structures on leaves [3]. In addition, the resemblance of larvae and adults of these pests can be subtle, with their general shapes being alike. This may could be the reason for error or misdiagnosis. Hugo et al. [21] underscore that there are damage symptoms and physical characteristics between the FAW and some stemborers, which can be difficult to distinguish. They cause similar types of damage to maize crops, including leaf feeding. Hence, Hugo et al. [21] conclude that understanding these similarities is crucial for developing effective integrated pest management strategies that can be applied across multiple pest species.

To sum up, one of the key findings of this chapter is that the FAW detection must be validated against Healthy and not non-FAW reports as there are good biological reasons for this behavior. This chapter has become evident by providing interpretation of the ineffectiveness of including Non-FAW category, as the data failed to meet the essential assumptions required for robust analysis. Furthermore, this finding explains why the *odds* of FAW when TN including total Healthy and Non-FAW cases (see Figure 3.9, purple curve) concluded that the FAW intensity was not significant in the maize filed.

Figure 4.10: Various pictures of Spotted Stemborer and violet stem borer, pictures from [2, 3, 4].

To the best of our knowledge, this is the first study to implement the meta-analysis techniques in a single study, yet across distinct locations. This analysis was a valuable study for multiple reasons. (i) All states applied the same methodology within an identical timeframe, providing the valid chance to apply the meta-analyses. (ii) The bivariate random effects logistic regression model allowed the research to consider the correlation between sensitivities and specificities, which is often ignored in the univariate approaches and observed in diagnostic test evaluations [171, 172]. (iii) The bivariate methodology can consider the observed variations of the Plantix app test accuracy between states. These variations may arise from a variety of factors. First, citizen scientists utilising the app may have different skills, education levels and smartphones with with differing camera quality. Second factor is the prevalence and severity of FAW infestations across different states which could also influence the accuracy of the diagnostic test. Indian states with heightened and more progressed FAW infestations may have displayed clearer symptoms and more easily identified FAW availability. Therefore, these states had higher sensitivities and lower specificities in comparison to states with lower levels

of infestation.

Moreover, our work is the first study to estimate the sensitivity and specificity of the Plantx app test, despite its use from 2016. Thus, it contributes to the literature on diagnostic test accuracy by providing reliable and robust estimates of the sensitivity and specificity of the Plantix app for detecting FAW in the maize. In addition, the same approach can be applied to other AI applications with citizen science. Additionally, the use of the MCMC algorithm in this analysis provided credible intervals for the sensitivity and specificity estimates, which capture the uncertainty associated with the point estimates [167]. In summary, the use of the bivariate random-effects logistic regression model combined with the MCMC algorithm allowed this study to obtain more precise and reliable estimates of the pooled sensitivity and specificity. Therefore, this approach provides a foundation for future research which could delve into novel methodological approaches for evaluating diagnostic tests in similar contexts.

However, while the bivariate random effects logistic regression model offered a robust approach to estimate the pooled sensitivity and specificity, it is essential to acknowledge the limitations of this study.

- First, the Plantix app test is the AI-based, hence, it may not be guaranteed to always provide correct answers. In other words, it was a chance to a tested case with high similarity to be wrong diagnosis, as well as it was a chance for a tested case with low similarity to be correct diagnosis.

- Second, the data used for the evaluation was based on the mixture models that did not entirely separate and overlap, potentially leading to challenges in precisely assessing the test's performance.

- Third, inaccuracies in the diagnostic process may be a result of the language barriers, blurry images or varying levels of users skill. The distance between the phone and the damaged symptom may lead to inaccurate diagnosis.

- Fourth, in our study, we initially assumed that the data collected from each state were independent and identically distributed (i.i.d.). However, it is possible that this assumption may not be valid in real-world scenarios, particularly when considering geographical proximity. Neighboring states may share similar environmental conditions, agricultural practices, and pest pressures, which could lead to correlated data. To address this limitation, future studies could consider models that

explicitly account for spatial dependence.

- Fifth, the research offers a comprehensive evaluation of the performance of the Plantix app in FAW detection using point estimates derived via the EM algorithm. While these estimates yield significant findings about the diagnostic accuracy of the Plantix app, they do not explicitly consider the uncertainty of the estimated parameters. Future work could address this limitation by employing statistical methods that includes the uncertainty of the estimated parameters., such as Bayesian estimation or resampling techniques.

Therefore, the estimation of the pooled sensitivity and specificity may not be completely accurate and is subject to limitations [171]. Hence, the results should be interpreted with caution, considering the potential limitations associated with the data sources and unmeasured confounders. Thus, the results of the study should be considered as estimates rather than definitive values.

Chapter 5

Mapping the prevalence of Fall Armyworm at state-level in India, by considering an AI imperfect diagnostic test

5.1 Introduction

In the realm of epidemiological studies in invasive pests, the accurate estimation of true prevalence within a population is crucial. This is because the true pest prevalence estimation provides numerical value of the infestation and outbreak across space and time. Such estimation helps us to understand the pest impact on crop health, productivity and the potential yield losses. It also aids in predicting the potential risks to human health. In addition, the prevalence estimation level over space and time facilitates understanding the dynamics of pest transmission and strategies to control and manage pest distribution [173]. Implementation of these strategies is achieved through prioritising, applying efficient resources to areas with higher pest pressure. The control strategy can include the frequent presence of pest monitoring specialists, providing farmers with suitable pesticides or offering monitoring tools such as satellites, drones or pest traps. Therefore, estimation of the true prevalence can support the establishment of a timeline and plan for monitoring and surveillance of crops, preventing pest invasion as a result.

5.1.1 True prevalence definition and estimation

Within the realm of statistics and research, true prevalence is the proportion of individuals in a population with a specific characteristic (specific disease, pest or condition) at a particular period of time [174, 175]. This metric is estimated by using a tested sample as a representation of the population. Hence, the true prevalence is calculated by dividing the number of individuals with the characteristic of interest by the total number of individuals in the sample [175]. In epidemiology and public health, a common method to investigate the presence of disease or pest in a distinct population is through applying diagnostic tests [176, 175]. Subsequently, the true prevalence in a perfect diagnostic test is the number of individuals who tested positive, divided by the total sample size [176]. However, a perfect diagnostic test is uncommon. Accordingly, if the diagnostic test is imperfect, the true prevalence estimation introduces bias [176, 175].

Therefore, to overcome this limitation, the concept of an apparent prevalence is introduced. An apparent prevalence is the proportion of individuals in a defined sample that test positive on a diagnostic test [176, 175]. Hence, the true prevalence and the apparent prevalence are equal when the diagnostic test is deemed perfect; otherwise, they differ [176, 175]. Thus, the true prevalence can be estimated from the apparent prevalence. However, the accuracy of apparent prevalence is directly impacted by the parameters of a diagnostic test; sensitivity (Se) and specificity (Sp) [175]. An adequate sample size and a random sample should also be considered in order to reach the accuracy of prevalence estimation. Table 5.1 presents the relationship between true prevalence, apparent prevalence and the parameters of a diagnostic test, where the detailed explanation can be found in subsection 5.2.2.

Table 5.1: The relationship between true prevalence (TPr), apparent prevalence (AP) and diagnostic test parameters (Se and Sp).

Diagnostic test	Has characteristic of interest	Does not have characteristic of interest	Total
positive	$TPr \times Se$	$(1 - TPr) \times (1 - Sp)$	AP
negative	$TPr \times (1 - Se)$	$(1 - TPr) \times Sp$	1- AP

Classic estimator and limitation

The Rogan-Gladen estimator, RGE (1978), incorporates both apparent prevalence and the test's characteristics (Se and Sp) to provide an unbiased and more accurate estimation of the true prevalence [177]. It is a classic estimator of true prevalence when the

diagnostic test is imperfect, and it is a well-known method in epidemiology [178]. The advantage of this estimator is its simplicity, where it applies a simple formula to calculate the true prevalence [178]. However, the estimator requires that both Se and Sp should be fixed and known [178, 175]. This assumption may lead to biased prevalence estimates. This is because it may not be practical in real-world applications [178, 175]. In addition, due to the nature of the formula, the true prevalence estimation value may fall outside the acceptable range, which should be between zero and one [178, 164]. Consequently, researchers should be cautious when using this estimator.

Enhancing prevalence estimation with MCMC

To address the limitations of the Rogan-Gladen estimator (RGE), the Bayesian statistics and the MCMC methods can be applied within the RGE. The MCMC methods enable researchers to capture the variability and uncertainty of the estimated parameters. As a result, they enhance the robustness and reliability of the true prevalence estimates. Moreover, the true prevalence parameter can be restricted to be within the interval $[0, 1]$ [178]. This can be achieved by selecting the uniform prior distribution on the interval $[0, 1]$ for the true prevalence parameter. This assumption is commonly used in studies, including the research conducted by Fischer et al. (2023) [104] and Flor et al. (2020) [164].

Furthermore, different factors in real-world settings such as environmental conditions or user behavior, can impact the characteristics of a diagnostic test (Se and Sp) [104]. Additionally, the random sampling process can be considered as another factor that influences these characteristics as well [164]. In addition, generating samples from various populations and subgroups can lead to heterogeneity in these metrics [104]. Thus, these factors have the potential to alter the Se and Sp rather than being fixed. Therefore, Se and Sp are often addressed as random variables to mitigate bias and adjust misclassification [164]. However, as a result, the number of unknown parameters in the RGE raises to be three which are Se, Sp and true prevalence [178]. Hence, the RGE becomes an over-parameterized equation.

Consequently, this complexity of the over-parameterized problem can be handled with the MCMC approach. Moreover, applying the Bayesian approach allows the researcher to capture the variability and uncertainty associated with these parameters (Se and Sp) [164]. Thus, prevalence estimation research has increasingly modelled them as

prior probability distributions [178, 164, 104]. To sum up, the Rogan-Gladen estimator within the Bayesian framework enhances the precision and reliability of prevalence estimates, by including the prior knowledge of true prevalence, sensitivity and specificity parameters.

Bayesian meta-analysis in true prevalence estimation

In Bayesian meta-analysis, the Rogan-Gladen estimator (RGE) is also a powerful tool to estimate the true prevalence. This approach can be extended to analyse data from multiple independent sources which may be conducted across diverse regions, as illustrated by Fischer et al. (2023) [104]. The Fischer et al. method [104] depends on a single diagnostic test across multiple regions. Therefore, the implementation of this technique promotes measurement consistency across regions, which can reduce bias resulting. In addition, using the meta-analysis concept by pooling data from several sources increases the overall sample size. This indicates that the findings mitigate sampling errors, leading to more stable and reliable estimations. Further, applying the meta-analysis rather than analysing individual studies separately, makes the comparisons of prevalence estimates between different geographic areas straightforward and easy. Furthermore, it reduces the possible bias due to diverse methodologies in distinct geographical regions.

5.1.2 Statistical analysis of FAW prevalence

A range of studies have demonstrated that almost the entire country of India provides suitable conditions for the FAW's persistent presence. They also confirmed the presence of FAW in various Indian states [179, 35, 136, 180, 104]. However, the statistical analysis of the true prevalence of FAW in Indian states have not yet been fully carried out. Therefore, a critical gap exists in our knowledge regarding the true prevalence of FAW. Although estimating true prevalence for non-native insects such as FAW is uncommon, statistically valuable knowledge can be gained from other disciplines.

One of the simplest and most straightforward statistical methods to estimate true prevalence is by using the 2x2 diagnostic table. This is because the proportion of the interesting characteristics is well-known and uncomplicated. Nevertheless, commonly, the gold standard list, or a reliable classification method, as we discussed in Chapters 3 and 4, should be available to construct the table [181]. Yet, even in the absence of a diagnostic table, other statistical techniques can be employed in order to estimate the

true prevalence. Currently, the recommended technique is Bayesian modelling [182].

Several studies estimating true prevalence have been conducted using multiple diagnostic tests. Combining several diagnostic tests can reduce challenges with prevalence estimation. Accordingly, the impact of bias that may be associated with a single test can be understood and minimised. Further, it enhances the accuracy of the estimation model where it offers a comprehensive analysis. These advantages were confirmed by Speybroeck et al. (2023) [183] in their study. They estimated the true prevalence of malaria infection in Peru, Vietnam, and Cambodia. They used a Bayesian framework with three diagnostic tests when no gold standard diagnostic test was available [183].

Moreover, based on two independent diagnostic tests, Sahlu and Whittaker conducted a study to estimate the true prevalence of the COVID-19 in Maryland [184]. The Bayesian framework was applied separately to each test. The incorporation of the Bayesian approach allowed to control misclassification errors by considering the sensitivity and specificity in the model. Hence, a binomial distribution was assumed to be the likelihood function. Its probability parameter of testing positive was estimated using RGE. Moreover, the model included prior probability distributions for both sensitivity and specificity. Therefore, this method emphasises that to improve the true prevalence accuracy regardless of the diagnostic test used, the adjusting misclassification errors should be included. This underscores the importance of applying the Bayesian methodology for correcting misclassification errors.

One such study that was based on a single imperfect diagnostic test, was addressed by Fischer et al. [104]. It utilised statistical techniques to estimate the true prevalence of major depressive disorder in Europe. Further, the approach was adopted in 27 European countries where the data source was based on the imperfect diagnostic accuracy of the PHQ-8 screening tool. Therefore, to account for the variability in diagnostic accuracy across different countries and provide more precise and reliable prevalence estimates, a Bayesian framework within meta-analysis was considered. In addition, this approach applied the latent class model to classify individuals into hidden categories, defining their health status as either a major depressive disorder, or not. By utilising these methodologies and adapting them to the specific challenges of FAW prevalence estimation, researchers can develop a more accurate estimation of FAW's true prevalence in Indian state.

5.1.3 Objective

The objective of this study was to estimate the true prevalence of Fall Armyworm (FAW) in maize crop over space and time. Subsequently, the geographical scale unit in the India was the Indian state, including states where the FAW is documented in the study data. Further, the timeframe was from the first day of 2018 to the end of 2019 and covered the known maize crops seasons in India, which are Kharif and Rabi. Hence, this is the epidemic phase of the FAW in India. Therefore, this study helped to understand the burden of FAW infestation during this critical period. To achieve this objective, four different statistical methods were used to find a more accurate estimation.

5.2 Materials and methods

5.2.1 Data description

The data analysed in this chapter was extracted using the exact methodology that was expounded upon in the data description subsection in the Chapter 4. A unique addition in this chapter is that the extraction of data for each state was based on the main seasonal timeframes of maize crops, Rabi and Kharif. Accordingly, there was a presentation of four 2×2 Tables shown for each state. The precise period for each season is detailed in Chapter 3, Kharif 2018 commenced on 10/6/2018 and concluded on 3/10/2018, Rabi 2018 occurred between 10/10/2018 and 28/3/2019, Kharif 2019 extended from 10/6/2019 to 3/10/2019 and Rabi 2019 spanned from 10/10/2018 to 31/12/2019.

5.2.2 Estimating true prevalence

This research estimated the true prevalence of the FAW in maize crops through four different methods, taking into account both temporal and spatial aspects. In terms of time, the analysis depends on the season of planting and harvesting timeframes of the maize crop. Spatially, the focus was on individual states, and the reasons for choosing states as the unit of spatial scale are explained in Chapter 4.

The rationale for employing four diverse methods was based on the understanding that while all methods might be statistically valid, there was a possibility that they would yield different findings. Hence, the determining of the most accurate approach and the most precise results were crucial. Consequently, this was achieved by comparing

with our previous result (*odds* plots in Chapter 3) and with a real-world maize field. Additionally, further reasons for adopting multiple methods were that due to the acknowledge imperfection of Plantix app data and the lack of a gold standard. Therefore, several methods were applied in order to be more cautious and to achieve more reliable results. Furthermore, a third rationale was to provide a valuable and comprehensive resource for researchers by presenting multiple methods in an individual document, enhancing their knowledge and critical thinking.

The following explanation presents each method, extensively describing its approach and assumptions. In addition, the methods are arranged in a sequential manner, starting from the easiest, most fundamental, and uncomplicated ones and progressing towards the more complex ones. Prior to detailing the methods, this study contributions are outlined below, focusing on the key improvements and differences from existing research:

(i) In Method-1, this study assumes that each state should satisfy a non-negative Yoadon index in order to estimate true prevalence. To the best of our knowledge, this assumption not been taken into consideration in the previous research.

(ii) Method-2 and Method-3 combine multiple statistical techniques, including the meta-analysis with the bivariate relationship between Se and Sp, Bayesian inference within MCMC, and traditional inference within the Rogan-Gladen Estimator (RGE) formula. This approach is applied to enhance the accuracy and reliability of prevalence estimates. The initial part is to estimate Se and Sp, using the the bivariate relationship between Se and Sp and Bayesian inference within MCMC statistical techniques. Next, the estimated Se and Sp are substituted in the RGE formula in order to estimate the true prevalence. Hence, these methods provide a comprehensive and powerful approach to statistical analysis. The meta-analysis which pools data from various states, can provide more precise estimates and improve the reliability of results.

(iii) Method-4 is similar to the model published by Fischer et al. [104]. However, the method applied here has additional likelihood functions and prior distributions, including priors for each parameter within it (hierarchical prior). Therefore, this method (Method-4) can lead to increase informative and accurate estimations and has the potential to reduce uncertainty. Further, having several likelihood functions means that the model may be built based on extra variables in the data, which enhances the estimation precision. In addition, with the hierarchical prior, the model's parameters can be represented in a greater flexibility and subtle way.

Method-1

The estimate of the true prevalence (TPr) of FAW in the maize crops in each state within a season, is easily computed by the following formula:

$$\text{True prevalence} = TPr_i = \frac{TP_i + FN_i}{n_i}. \tag{5.1}$$

where i denotes each distinct state. This formula computes the proportion of maize crops which has the FAW among the total number of the given cases. To consider the resulting value of the FAW prevalence in a particular state, at a given season, two assumptions were taken into account.

$$\begin{aligned} &\text{C0:} && \text{sample size} && > 30 \\ &\text{C1:} && \left(\text{Se}_{i,\,FAW} + \text{Sp}_{\overline{FAW}} - 1\right) && > 0 \end{aligned}$$

C0 plays a vital role in mitigating the risk of bias and representing the entire study population. C1 is a fundamental criterion which it must be applied to any diagnostic test. If C1 is not upheld, the test detects the interested characteristic by chance alone. Further, C1 is equivalent to the Youden index (J). For more details, see Chapter 4. In addition, to calculate the confidence interval (CI) for each estimated prevalence, the Wilson estimator (which ensures the results stay within the range of 0 to 1) was applied using the binconf function from the Hmisc package in R.

Method-2

Implementing an apparent prevalence (AP), Se_{FAW}, and $\text{Sp}_{\overline{FAW}}$ to estimate the true prevalence (TPr) in one function can be an alternative method rather than Method-1. This cosideration is important in the presence of diagnostic test imperfection. If the diagnostic test is imperfect, AP is subject to the bias due to have FP or FN events [185, 176]. Realising this bias underscores the importance of considering the test's imperfection by deriving the true prevalence from AP, Se, and Sp values.

Therefore, by including both Se and Sp into the estimation model, a more accurate representation of the true prevalence (TPr) can be achieved. In this instance, Se correctly accounts for the true positives, while Sp effectively excludes the false positives [185, 176]. Subsequently, AP, the sum of the probabilities of true positives (TP) and

false positives (FP), can be used to express the TPr:

$$\begin{aligned} AP &= p(TP + FP) \\ &= p(TP/D)\, p(D) + p(FP/ND)\, p(ND) \qquad (5.2) \\ &= Se \times TPr + (1 - Sp) \times (1 - TPr). \qquad (5.3) \end{aligned}$$

To solve for the true prevalence, TPr;

$$\begin{aligned} AP = (Se \times TPr) - & \\ & TPr + (1 - Sp) + (Sp \times TPr) \\ TPr - (Se \times TPr) - (Sp \times TPr) &\Longrightarrow -AP + (1 - Sp) \\ TPr \times (1 - Se - Sp) &\Longrightarrow -AP + (1 - Sp) \\ TPr \times (Se + Sp - 1) &\Longrightarrow AP - (1 - Sp) \\ TPr &\Longrightarrow \frac{AP - (1 - Sp)}{1 - Se - Sp} \\ &\Longrightarrow \frac{AP + Sp - 1}{Se + Sp - 1} \\ TPr &= \frac{AP + (Sp - 1)}{J}. \end{aligned} \qquad (5.4)$$

Following this, the values found in Chapter 4 for the point estimate (mean) Se and Sp of the Plantix app accuracy can be entered into the formula (5.4) to yield $TPr = \frac{AP+(0.77-1)}{0.49+0.77-1}$. Subsequently, the estimated true prevalences displayed different results among states due to the differences in the observed AP values. However, it is important to note that the results obtained from formula (5.3) may exceed one or be less than zero. Hence, to guarantee that the true prevalence estimations falls within the proper range of zero to one, the following assumptions and conditions must be satisfied for each state:

$$\begin{aligned} &\text{C0:} \quad \text{sample size} \quad > 30 \\ &\text{C1:} \quad (1 - Sp) \quad < Se \\ &\text{C2:} \quad (1 - Sp) \quad \leq AP \\ &\text{C3:} \quad AP \quad \leq Se \end{aligned}$$

In this case, C1 is equivalent to the Youden index (J), hence, it can be represented as Se +Sp - 1 > 0. Inability to meet the C2 yields a negative value, $TPr < 0$. Likewise,

failing to achieve C3 leads to $TPr > 1$. To sum up, Method-2 is an amalgamation of the meta-analysis, Bayesian inference and traditional inference to derive estimates of the true prevalence. To calculate the credibal interval (CrI) for each state in this method, the lower and upper credible interval values of the pooled estimate (mean) Se and Sp of the Plantix app accuracy found in Chapter 4 can be entered into the formula (5.4).

Method-3

There are two steps in this method and they are applied in each season. The first step employes the Bayesian framework. The purpose of this is to estimate the pooled Se and Sp of the Plantix app test. The bivariate generalized linear mixed effects model of the meta analysis within the MCMC framework is used. This step is identical to the description in Chapter 4. Moreover, the same assumptions (C0 and C1) are considered when selecting the states and before implementing the MCMC. Afterwards, the seasonal point estimate (mean) of the Se and Sp values from the first step is utilised in the second step. It is important to note that these values could differ between seasons. After that, to compute the TPr for each state within season, the frequentist statistics formula (5.4) is applied. In addition, the above-mentioned assumptions C2 and C3 are also taken into account to guarantee relaible estimations. After that, to calculate the credibal interval (CrI) for each state in this method, the lower and upper credible interval values of the pooled estimate (mean) Se and Sp of the Plantix app accuracy found in each seaseon, can be entered into the formula (5.4).

To sum up, the integration of the meta-analysis, Bayesian inference and traditional inference to estimate the diagnostic test accuracy and true prevalence, can capture complex relationships between prevalence, sensitivity, and specificity. As a result, it can enhance the estimation accurcy. Moreover, the consideration of the temporal trends and spatial variations in determinants of true prevalence offers more precise estimation. Therefore, this method contributes an advanced and thorough approach to provide a more comprehensive understanding of the true prevalence.

Method-4:

This method is a Bayesian hierarchical model and was applied separately in each season. The model assumed binomial distributions as the likelihood functions for the counts of

true positives TP_i, true negatives TN_i, and for the total positive tests r_i $(TP_i + FP_i)$. The binomial distribution was commonly used as likelihood for r_i [185, 176]. Hence, the Se, Sp and AP_i were the probability of success parameters in each state.

$$TP_i \sim Binomial(n_{i,\,FAW}, Se_i).$$
$$TN_i \sim Binomial(n_{i,\,\overline{FAW}}, Sp_i).$$
$$r_i \sim Binomial(n_i, AP_i).$$

n_i is the total sample size in each state within a season. $n_{i,\,FAW}$ is the total number of $(TP_i + FN_i)$ in each state i , while $n_{i,\,\overline{FAW}}$ is the total number of $(TN_i + FP_i)$ in each state i. The C0 and C1 should be also considered for the selected states in each season. To construct the marginal distribution for the true prevalence for each state, the MCMC steps incorporate the RGE:

$$AP_i = Se_i \times TPr_i + (1 - Sp_i) \times (1 - TPr_i).$$

Moving on to the prior distributions for each of parameters of interest, the non-informative prior was used to represent the prevalence parameter in each state and within each season. A beta prior was used for the prevalence TPr_i with a uniform distribution, $beta(1, 1)$. This beta distribution is a flexible distribution that is often used to model probabilities or proportions, as defined on the interval $[0, 1]$. Additionally, to account for the potential variability, and correlations between Se and Sp, a multivariate normal prior distribution for the logits of sensitivity and specificity with mean vector $\boldsymbol{\beta}$ and covariance matrix Ω was applied.

$$TPr_i \sim beta(1, 1).$$

$$logit \begin{bmatrix} Se_i \\ Sp_i \end{bmatrix} \sim MVN \left[\boldsymbol{\beta} = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}, \quad \Omega \right], \Omega = \begin{bmatrix} \tau_1^2 & \tau_1\tau_2\rho \\ \tau_1\tau_2\rho & \tau_2^2 \end{bmatrix}.$$

Further, prior distributions for the parameters within the multivariate normal distribution were included. For the mean vector β;

$$\boldsymbol{\beta} = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} \sim MVN(\mu, prec) = MVN \left[\boldsymbol{\mu} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \Sigma = \begin{bmatrix} 0.1 & 0 \\ 0 & 0.1 \end{bmatrix} \right].$$

Additionally, the prior distribution of Ω matrix was

$$\mathbf{\Omega}^{-1} \sim \text{ Wish } \left(S = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, df = 2 \right).$$

where Wish indicates an inverse-Wishart distribution. In this situation, Ω is to model the variability in sensitivity and specificity parameters among studies**.** The hierarchical priors enable this study to achieve flexible modeling of the parameters. Further, the model takes into consideration the uncertainty in the parameters and incorporated prior knowledge about the relationships between sensitivity and specificity values across studies. In addition, the model does not explicitly specify a prior distribution for AP_i; instead, it is indirectly shaped by the priors of Se, Sp and TPr_i. This is because AP_i is a function, or dependent variable, of these three parameters.

Finally, after specifying the model, the next step involves running the MCMC algorithm to estimate the posterior samples of the parameters, particualarly the marginal posterior of the prevalence. Stan, a probabilistic programming language, [186] is used for MCMC sampling, with three chains, 10,000 iterations, and 2000 warm-up iterations. The trace plots, R-hat values, effective sample size (ESS), and autocorrelation plots (ACF) are examined to assess model convergence. Further, the MCMC algorithm is run twice once with the original sample size for each state, and again with 10,000 duplicate observations for each included data variable in the model. This carries out in order to maintain consistency, and reduce uncertinity as the variance decreased and indicated a negative relationship with the sample size.

5.3 Results

This section presents the prevalence of FAW at state level in India over four different seasons; the maize crops periods from 2018 to the end of 2019. This timeframe is the epidemic phase of the FAW infestation in India. Therefore, to gain a comprehensive understanding of the prevalence, establishment, and distribution of the FAW during the period of the outbreak, four different statistical methods were applied. Hence, by using a variety of methods, the most accurate representation of the prevalence rates could be identified. Accordingly, this section is divided into four parts, where each represented a specific season. Within each season, the findings from four different statistical methods are presented in details:

1. Prevalence of FAW during Rabi 2018

2. Prevalence of FAW within Kharif 2018

3. Prevalence of FAW in Rabi 2019

4. Prevalence of FAW across Kharif 2019

5.3.1 Estimation of the Prevalence of FAW at state-level during Kharif 2018

The following Table presents the prevalence estimation in Kharif 2018 across Indian states. The states in Table **??** are arranged in descending order based on the total sample size (total).

Table 5.2: FAW prevalence at Indian-state level with total sample sizes greater than 30 during Kharif 2018 (C0), using four different prevalence statistical methods.TP is the number of true positives, and FN is the number of false negatives. P is the total positives, calculated as TP plus false positives (FP). AP is the apparent prevalence. Total refers to the total maize tested in the state. C1, C2 ($AP - (1 - Sp) \geq 0$), and C3 ($Se - AP \geq 0$) are the criteria for selecting states. TPr1 to TPr4 are the estimated true prevalence based on Methods 1 to 4. n=10000 means the sample size for Method-4 is multiplied by 10000. The (–) indicates that the true prevalence estimations of state are invalid due to not meeting the criteria.

						Method-1		Method-2			Method-3			Method-4
State	TP	FN	P	Total	AP	C1	TPr1 (95% CI)	C2	C3	TPr2 (95% CrI)	C2	C3	TPr3 (95% CrI)	TPr4, n=10,000 (95% CrI)
Maharashtra	81	265	149	877	0.17	0.10	0.39 (0.36,0.43)	-0.06	0.32	-	-0.01	0.21	-	0.39 (0.39, 0.4)
Uttar Pradesh	30	164	100	554	0.18	-0.04	-	-0.05	0.31	-	0	0.2	-	-
Madhya Pradesh	28	139	81	491	0.16	0.01	0.34 (0.30,0.38)	-0.07	0.33	-	-0.02	0.22	-	0.34 (0.22,.38)
Rajasthan	19	123	56	394	0.14	-0.02	-	-0.09	0.35	-	-0.04	0.24	-	-
Telangana	100	89	131	381	0.34	0.37	0.50 (0.45,.55)	0.11	0.15	0.41 (0.39,.52)	0.16	0.04	0.81 (0,53,1)	0.50 (0.49,0.5)
Andhra Pradesh	58	66	80	230	0.35	0.26	0.54 (0.47,0.60)	0.12	0.14	0.45 (0.40,.54)	0.17	0.03	0.87 (0,55,1)	0.54 (0.54,.54)
Karnataka	66	34	84	187	0.45	0.45	0.53 (.46, 0.60)	0.22	0.04	0.83 (0.67,1)	0.27	-0.07	-	0.53 (0.53,.54)
Haryana	10	51	38	164	0.23	-0.11	-	0	0.26	0 (0,.1)	0.05	0.15	0.24 (0,24,1)	-
Gujarat	27	38	51	160	0.32	0.17	0.41 (0.33,0.48)	0.09	0.17	0.34 (0.33,.34)	0.14	0.06	0.71 (0.47,1)	0.41 (0.40,.41)
Bihar	2	38	15	126	0.12	-0.10	-	-0.11	0.37	-	-0.06	0.26	-	-
Chhattisgarh	10	36	24	113	0.21	0.01	0.41 (0.32, 0.5)	-0.02	0.28	-	0.03	0.17	0.14 (0,.19)	0.41 (0.28,.45)

					Method-1		Method-2			Method-3			Method-4
State	TP	FN	P	Total	AP	C1	TPr1 (95% CI)	C2	C3	TPr2 (95% CrI)	C2	C3	TPr3 (95% CrI)	TPr4, n=10,000 (95% CrI)
Punjab	10	27	27	93	0.29	-0.03	-	0.06	0.20	0.22 (0,14,.25)	0.11	0.09	0.55 (0.40,1)	-
Jharkhand	8	23	20	85	0.24	0.04	0.36 (0.27,.47)	0.01	0.25	0.03 (0,0.11)	0.06	0.14	0.29 (0.28,1)	0.36 (0.33,.38)
Tamil Nadu	21	19	27	79	0.34	0.37	0.51 (0.40, .61)	0.11	0.15	0.41 (0.39,50)	0.16	0.04	0.81 (0.53,1)	0.51 (0.50,.51)
Himachal Pradesh	0	5	4	45	0.09	-0.10	-	-0.14	0.4	-	-0.09	0.29	-	-
West Bengal	5	9	14	35	0.40	-0.07	-	0.17	0.09	0.64 (0.54,.91)	0.22	-0.02	-	-

Method-1

We employed Method-1 to estimate the prevalence ($TPr1$) of FAW infestation at the state level across India. The prevalence was calculated using the formula TPr. For example, in Andhra Pradesh, among 230 maize cases, 58 were classified as TP notifications, while 66 were erroneously diagnosed as FN. Hence, by applying the prevalence formula (5.1), the prevalence of FAW in Andhra Pradesh was estimated as $TPr1 = \frac{(58+66)}{230} = 0.54$ equivalent to 54 percent.

On the map shown in Figure 5.1, the prevalence estimated values are shown, where the darker colours reflect the higher level of prevalence. As may be observed, the FAW prevalence level across all studied states varied between 0.34 and 0.54. Andhra Pradesh (54 percent prevalence), Karnataka (53 percent), Tamil Nadu (51 percent) and Telangana (50 percent) in the southeast of India reocrded the highest level of infestation compared to other regions. During this season, Chhattisgarh (41 percent) which is located in North India, bordering Telangana and Andhra Pradesh to their north, also exhibited a high level of infestation.

Further, Tamil Nadu and Andhra Pradesh are coastal states. It is possible that the FAW entered India via the coastline and the initial appearance was therefore in the southern states. Moreover, Gujarat which is the most westerly state in India, had a prevalence of 41 percent during this season, making it among the highest rates observed. This could be due to the state having coastal boundaries on all but one sides, to the east. It is surrounded by the Arabian Sea on the west and southwest, the Gulf of Khambhat on the south, and the Gulf of Kutch in the northwest.

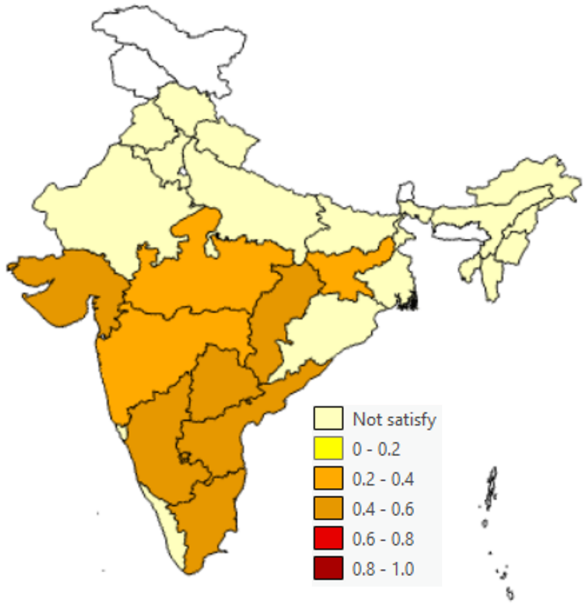


Figure 5.1: Mapping the prevalence of FAW infestation in the maize crop during Kharif 2018 at state level using Method-1. A higher prevalence is associated with darker colours, while the lightest yellow indicates unsatisfied assumptions. No colour suggests that the area is not under study.

Method-2

In addition, Method-2 was applied to estimate FAW prevalence ($TPr2$) during Kharif 2018 in each of the states under study. For example, the estimated prevalence in Karnataka was $TPr2 = \frac{AP-(1-Sp)}{\text{Se}+Sp-1} = \frac{0.45-(1-0.77)}{0.49+0.77-1} = 0.83$ or 83 percent. The same formula was applied to each state, where the only differences between states was the AP rate. Seven states Chhattisgarh, Uttar Pradesh, Maharashtra, Madhya Pradesh, Rajasthan, Bihar and Himachal Pradesh have negative $TPr2$ values. This occurred because these states failed to satisfy condition-two (C2), where C2 is specified as $1 - Sp \leq AP$, or alternatively, $(1 - Sp) - AP \geq 0$. Hence, the AP values in these states were lower than FPR $(1 - Sp)$. As a result, Method-2 when applied in this season, the method estimated the FAW prevalence for nine of the fifteen states. The estimates ranged between 0 percent in Haryana to 83 percent in Karnataka.

Furthermore, Method-2 agreed with Method-1, suggesting that FAW originated in the coastal states of southeast India. Based on Method-2 findings, Karnataka had the highest prevalence at 83 percent. An explanations for this could be that Karnataka was the first state in India to officially report cases of FAW. However, the 83 percent value may potentially be an overestimation at the start of the FAW season in India.

Moreover, 64 percent of maize crops were infested with FAW in West Bengal, which is located on the eastern coast. This may also be overestimated and less certain, because its rate was higher than the costal south states.

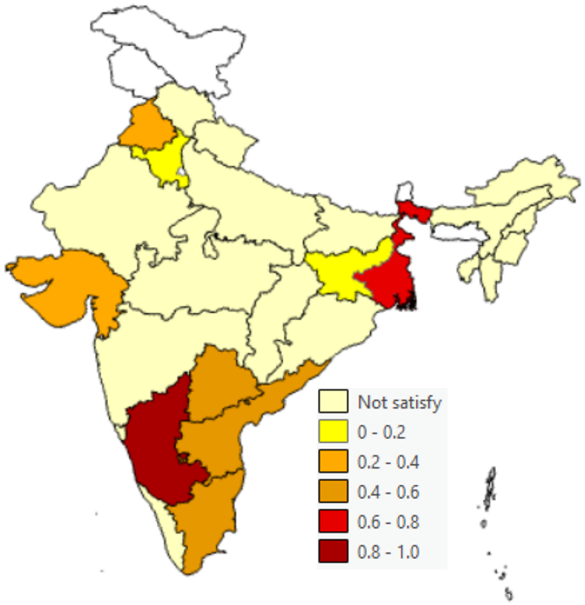


Figure 5.2: Mapping the prevalence of FAW infestation in the maize crop during Kharif 2018 at state level, using Method-2. A higher prevalence is associated with darker colours, while the lightest yellow indicates unsatisfied assumptions. No colour indicates that the area is not under study.

Method-3

As part of this method, the MCMC algorithm was employed to estimate the pooled Se and Sp for Kahrif 2018. To ensure the reliability of posterior estimates, the convergence of the three MCMC chains was assessed. Accordingly, four various diagnostic tools were evaluated to test convergence, after discarding a burn-in period of 3,000 iterations, a total of 10,000 iterations were conducted.. The trace plots of both parameters (Se and Sp) indicate an overlapping of the three chains and that they are well-mixed. Further, both parameters had single peak distributions as can be seen in the density plots. This is an indication that the model was able to estimate the parameter space. In addition, the ACF plots of both parameters show a sharp decrease in autocorrelation values and fluctuate around zero, which in turn, indicates a well-mixed chain and that they effectively explore the posterior distribution, thus confirming convergence. Furthermore, the parameters have $\hat{R} = 1$. The ESS rates for both parameters were 100 percent which is

the highest possible value, indicating that all samples effectively contributed to estimate the parameter spaces as well as being well mixed and independent samples. Therefore, mixing well, overlapping chains, unimodel distributions, ACF values fluctuating near zero, $\hat{R} = 1$and the maximum rates of ESS provide a reliable and accurate estimation of the MCMC.

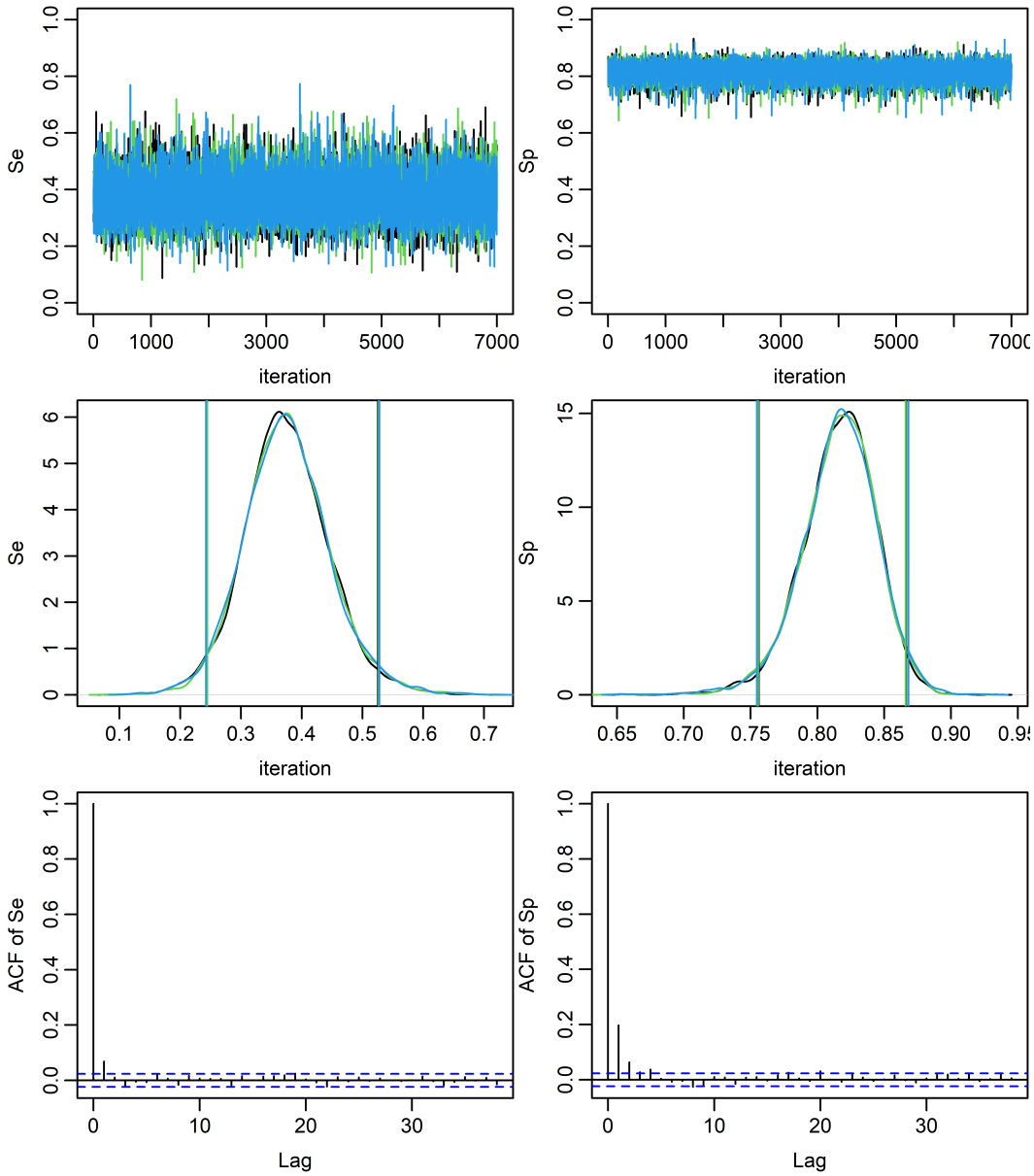


Figure 5.3: Assessing the convergence of three MCMC chains (post burn-in) for sensitivity and specificity during Kharif 2018, using trace plots in the first row, density plots in the second row, and ACF Plots in the last row.

Hence, the estimated pooled Se and Sp for states during Kharif season is shown in Table 5.5. By comparing the estimated Se and Sp for this season with the entire

period (Method-2), it can be observed that the Se is lower. On the other hand, the Sp is higher. This variability must not be overlooked. It may highlight that assessing sensitivity and specificity for each season may result in improved accuracy.

Table 5.5: Point estimates and credible intervals of MCMC summary results for sensitivity (Se) and specificity (Sp) parameters during Kharif 2018.

	Mean	95 CrI
Pooled Se	0.376	0.243, 0.526
Pooled Sp	0.816	0.756, 0.866

Following this, the implementation of the TPr3 formula was undertaken to estimate the FAW prevalence during the Kharif 2018 season. The estimated prevalence rates can be seen in Table **??**tab:Khari18prevalence under the column labeled TPr3. During this season, the TPr3 estimation was possible for just eight states. It was also verified that the outbreak had originated along the southern coast states. However, the estimations seem to be exaggerated. This is because it is not logical for the prevalence to be so excessively high in the first season. Further, the estimation exceeded that of the previous method for this season. One explanantion for this discrepancy may be attributed to the higher Sp value.

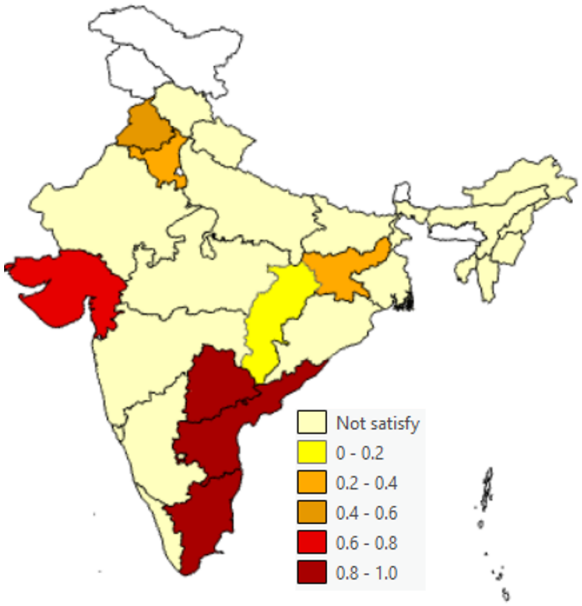


Figure 5.4: Mapping the prevalence of FAW infestation in the maize crop during Kharif 2018 at state level using Method-3. A higher prevalence is associated with darker colours, while the lightest yellow indicates unsatisfied assumptions. No colour indicates that the area is not under study.

Method-4

During this season, Method-4 could derive estimates for only nine of the states included in the study. Those included met the criteria of having a sample size greater than 30 (C0) and non-negative J value (C1). Following this, upon examining the MCMC with 10,000 iterations and three chains, along with a burn-in period of 2,000 iterations, it was noted that only two states Telangana and Karnataka exhibited signs of convergence (see in the Appendix Figure A.1).

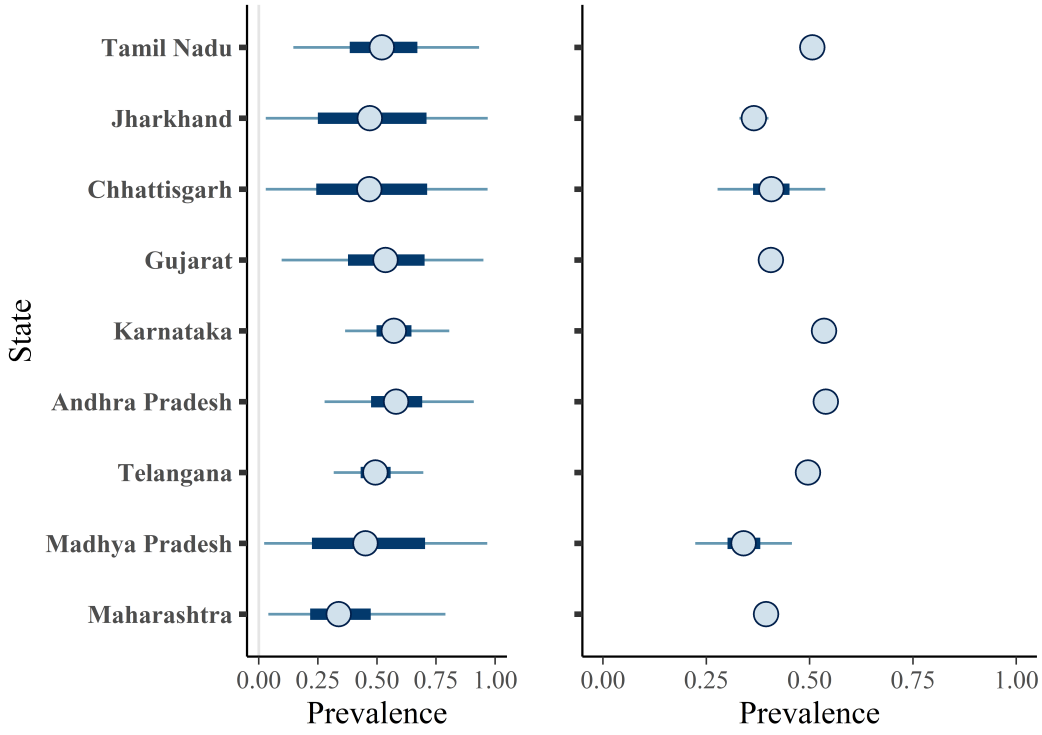


Figure 5.5: Forest plots for the mean posterior prevalence estimates and credible intervals (CrI) of the Indian states during Kharif 2018, where the bold line indicates 50% Crl, and the thin line indicates 95% CrI. The left plot represents forest plot for the original sample size for each state, while the right panel illustrates the forest plot for each state after scaling the sample size by a factor of 10,000.

As previously noted, the model is experiencing low convergence, likely due to the limited sample size. To address this, the proposed solution is to augment the obsewrations by multiplying each value of TP, TN, FP, and FN by a large constant (10,000) and then rerunning the analysis. However, it is important to note that, augmenting the sample size for each population has no effect on the point estimate. Nevertheless, it effectively diminishes uncertainty through lowering the variance within the population. This is because the reduction in variance is directly correlated to the augmentation in sample

size (n), with a mathematical relationship of $1/n$.

As a result, it can be seen in the right panel of Figure (5.5) that the forest plot demonstrates a decrease in uncertainty in all nine states. In addition, it can be seen that the trace plots in Figure A.2 in Appendiz suggest a visual evidence of convergence. To elaborate, the trace plots for the nine states were mixed effectively. Consequently, this procedure led to convergence for all states. Moreover, by comparison with Method-1, the estimated values were matched.

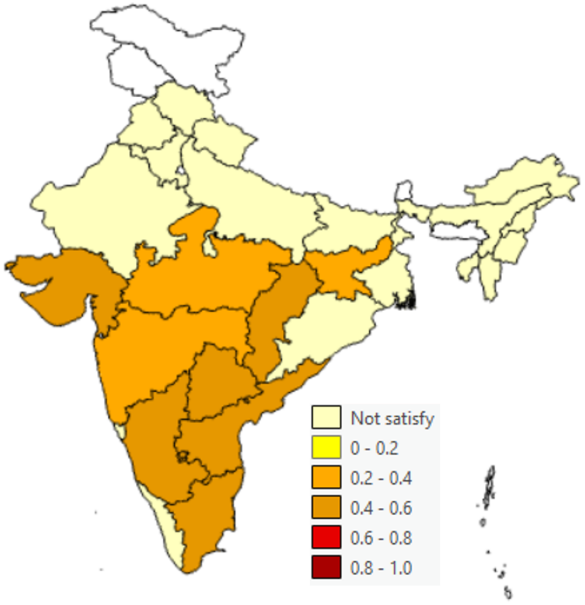


Figure 5.6: Mapping the prevalence of FAW infestation in the maize crop during Kharif 2018 at state level using Method-4 after scaling the sample size for each state by a factor of 10,000. A higher prevalence is associated with darker colours, while the lightest yellow indicates unsatisfied assumptions. No colour indicates that the area is not under study.

5.3.2 Estimation the prevalence of FAW at state-level during Rabi 2018

Table 5.6: FAW prevalence in Indian-state level during Rabi 2018, using four different prevalence statistical methods, where each state has total sample size no less than 30. TP is the number of true positives, and FN is the number of false negatives. P is the total positives, calculated as TP plus false positives (FP). AP is the apparent prevalence. Total refers to the total maize tested in the state. C1, C2, and C3 are the criteria for selecting states. TPr1 to TPr4 are the estimated true prevalence based on Methods 1 to 4. n=10,000 means the sample size for Method-4 is multiplied by 10,000. The (–) indicates that the true prevalence estimations of state are invalid due to not meeting the criteria.

						Method-1		Method-2			Method-3			Method-4	
State	TP	FN	P	Total	AP	C1	TPr1 (95 CI)	C2	C3	TPr2 (95 CrI)	C2	C3	TPr3 (95 CrI)	TPr4	TPr4, n=10,000 (95 CrI)
Andhra Pradesh	481	178	561	1017	0.55	0.51	0.65 (0.62,.68)	0.32	-0.06	-	0.38	-0.01	-	0.65	0.65 (0.56,.56)
Telangana	311	101	350	609	0.57	0.55	0.68 (0.64,.71)	0.34	-0.08		0.4	-0.03	-	0.68	0.68 (0.68,.68)
Maharashtra	197	103	234	529	0.44	0.50	0.57 (0.52,.61)	0.21	0.05	0.79 (.66,1)	0.27	0.1	0.72 (.52,1)	0.57	0.57 (.57,.57)
Karnataka	190	96	233	516	0.45	0.47	0.55 (0.51,0.6)	0.22	0.04	0.83 (.68,1)	0.28	0.09	0.75 (.54,1)	0.56	0.55 (.56,.56)
Tamil Nadu	139	36	163	254	0.64	0.49	0.69 (0.63,.74)	0.41	-0.15	-	0.47	-0.1	-	0.70	0.69 (.69,.69)
Bihar	7	40	17	245	0.07	0.10	0.19 (0.15,.25)	-0.16	0.42	-	-0.1	0.47	-	0.31	0.19 (.19,.2)
Gujarat	30	51	51	174	0.29	0.14	0.47 (0.39,.54)	0.06	0.2	0.22 (.16,.26)	0.12	0.25	0.32 (.3,.62)	0.51	0.47 (.46,.47)
Uttar Pradesh	7	62	21	152	0.14	-0.07	-	-0.09	0.35	-	-0.03	0.4	-	-	-
Madhya Pradesh	4	42	17	126	0.13	-0.07	-	-0.1	0.36	-	-0.04	0.41	-	-	-
West Bengal	6	26	20	95	0.21	-0.03	-	-0.02	0.28	-	0.04	0.33	0.10 (0,.17)	-	-
Rajasthan	4	30	16	92	0.17	-0.09	-	-0.06	0.32	-	0	0.37	-	-	-
Haryana	3	15	6	38	0.16	0.02	0.47 (0.32,.63)	-0.07	0.33	-	-0.01	0.38	-	0.48	0.47 (.38,.57)
Chhattisgarh	4	12	12	37	0.32	-0.13	-	0.09	0.17	0.34	0.15	0.22	0.4 (.34,1)	-	-
Punjab	2	16	4	35	0.11	-0.01	-	-0.12	0.38	-	-0.06	0.43	-	-	-

Method-1

Based on Method-1 and during Rabi 2018, the highest infestation prevalence was estimated at 69 percent in Tamil Nadu, while in Bihar, which is located in the northeast, displayed the lowest prevalence at 19 percent (see Figure 5.7). Further, there was a significant escalation in the FAW prevalence intensity in Tamil Nadu, Telangana (68 percent), and Andhra Pradesh (65 percent), when compared with the previous season and when appling the same method. Furthermore, a slight two percent rise in prevalence was observed in Karnataka. This season was the time at which India officially acknowledged the presence of FAW in the country, starting in Karnataka. The announcement may have prompted control and mitigation measures, that restriced the magnitude of the escalation in Karnataka.

Moreover, Maharashtra recorded a substantial increase in the prevalence, rising from 39 percent to 57 percent. This increase might be linked to the state's location, where Gujarat lies on its western border as well as Karnataka, Chhattisgarh, and Telangana on its southern and eastern boundaries. In the North India, Haryana and Bihar were found to have prevalence rates of 47 percent and 19 percent, respectively, while there were no prevalence recorded in these states during the previous season using this method. It is worth mentioning that Bihar cultivates more maize in the Rabi season than in the Kharif season in [40]. Further, this method failed to estimate the FAW prevalence in Chhattisgarh, Madhya Pradesh and Jharkhand during the same season, despite these states having recorded a prevalence during the Kharif season. Compared to Kharif, these states are anticipated to have planted fewer maize crops during this season.

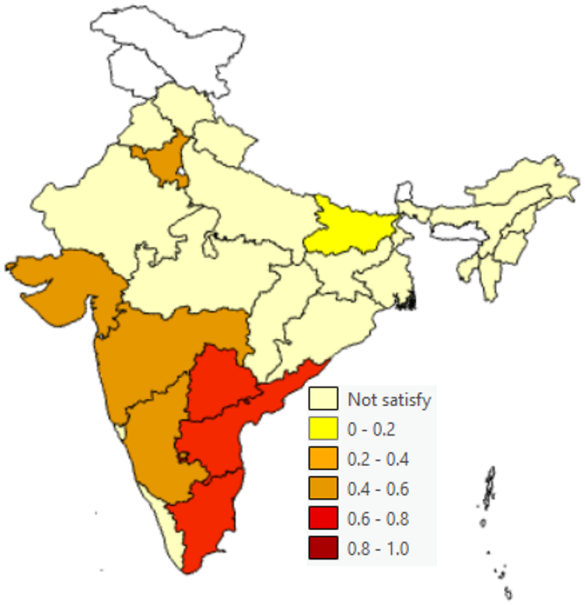


Figure 5.7: Mapping the prevalence of FAW infestation in the maize crop during Rabi 2018 at state level using Method-1. A higher prevalence is associated with darker colours, while the lightest yellow indicates unsatisfied assumptions. No colour indicates that the area does not included under this study.

Method-2

Based on this method, not all states satisfied C2 or C3, with some notable exceptions Karnataka, Maharashtra, Chhattisgarh and Gujarat. The prevalence estimation for these states were 79, 83, 22 and 34 percent, respectively. Compared to Method-1, this approach assigned a higher value to Karnataka, Maharashtra and Chhattisgarh. Figure 5.8 shows the FAW prevalence based on this method.

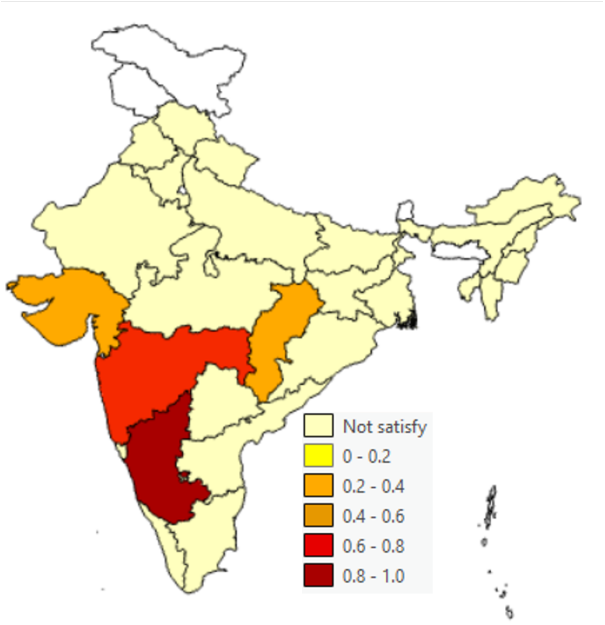


Figure 5.8: Mapping the prevalence of FAW infestation in the maize crop during Rabi 2018 at state level using Method-2. A higher prevalence is associated with darker colours, while the lightest yellow indicates unsatisfied assumptions. No colour indicates that the area is not under study.

Method-3

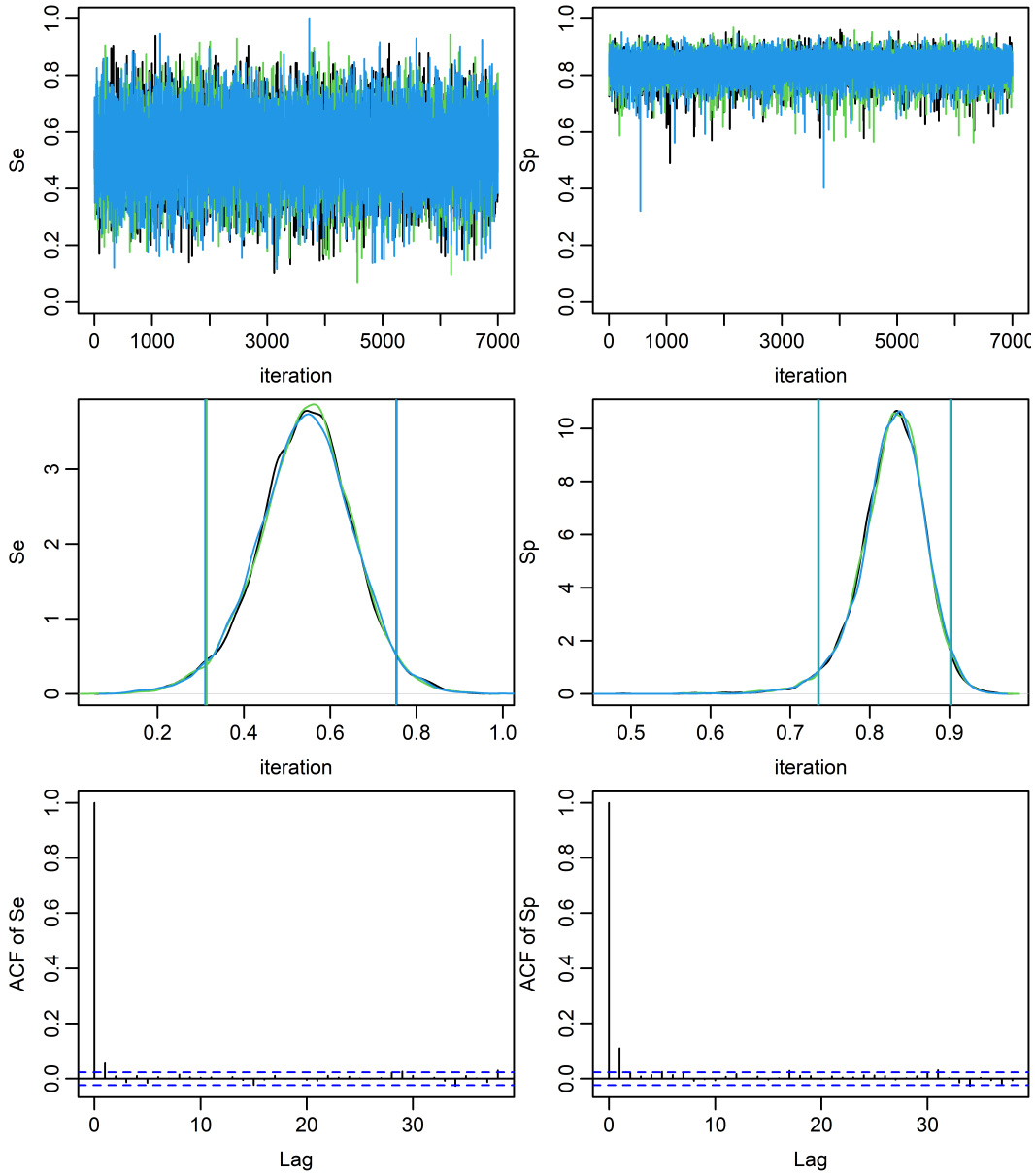


Figure 5.9: Assessing the convergence of three MCMC chains (post burn-in) for sensitivity and specificity during Rabi 2018, using trace plots in the first row, density plots in the second row, and ACF Plots in the last row.

Further, the MCMC was employed to estimate the pooled *Se* and *Sp* for Rabi 2018. To ensure the reliability of posterior estimates, the convergence of the three MCMC chains was assessed. After removing a 3000 burn-in period of 10,000 iterations, four various diagnostic tools were used to test convergence. The trace plots of both parameters (*Se* and *Sp*) display overlapping of the three chains, thus indicating that they are well-mixed. Although *Se* exhibits a wider variation, both parameters had single peak distributions

as can be seen in the density plots. This is an indication that the model was able to estimate the parameter space.

In addition, the ACF plots of both parameters show a rapid decline towards zero, which further indicates an effective exploration of the posterior distribution, thus confirming convergence. Furthermore, the parameters have $\hat{R} = 1$. Additionally, the ESS rates for both parameters were 100% which is the highest possible value. This indicates that all samples effectively contribute to estimating the parameter spaces as well as well mixed and independent samples. Hence, well mixed and operlapping chains, unimodel distributions, ACF values fluctuating near zero, $\hat{R} = 1$and the maximum rates of ESS, provide a reliable and accurate estimation of the MCMC. Therefore, Table 5.7 summarizes the pooled Se and Sp for states during Rabi 2018. To compare the estimated Se and Sp for this season with the whole period (Method-2), it can be noted that if rounded to one decimal place, the values are the same. Nevertheless, with additional decimal places, the values for the entire period were lower than those for this season.

Table 5.7: Point estimates and credible intervals of MCMC summary results for sensitivity (Se) and specificity (Sp) parameters during Rabi 2018.

	Mean	95 CI
Pooled Se	0.543	0.310, 0.754
Pooled Sp	0.828	0.734, 0.901

Moving forward, the study commenced with applying the TPr3 formula in order to estimate the FAW prevalence during the Rabi 2018 season. Table 5.6, within the column entitled TPr3, presents the estimated prevalence values. The TPr3 indicated valid values, which should be between 0 and 1, for Maharashtra (72 percent), Karnataka (75 percent), Gujarat (32 percent), West Bengal (10 percent) and Chhattisgarh (40 percent). From the map (see Figure 5.10), it is evident that Method-3 could not estimate the prevalence rates among the states along the southeast coastline. This is because the apparent prevalence exceeded the sensitivity in these states. Hence, the FP cases exceeded the TP cases. As a result, the true prevalence values were overestimation. To sum up, the formula failed to estimate the majority of states, which resulted in an inaccurate representation of the prevalence dynamics during this period.

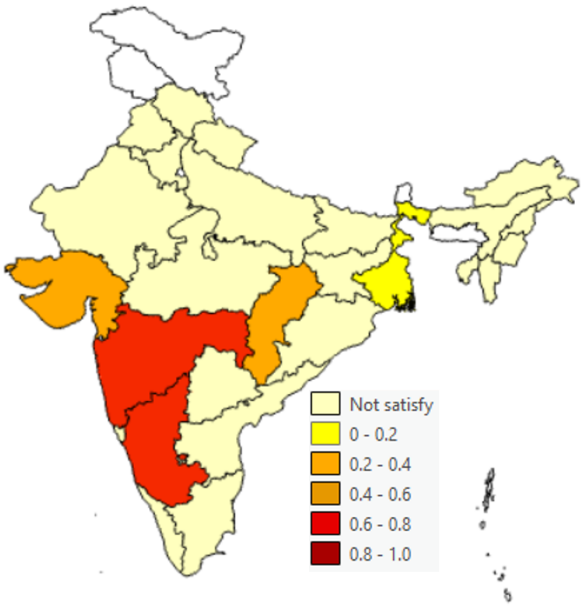


Figure 5.10: Mapping the prevalence of FAW infestation in the maize crop during Rabi 2018 at state level using Method-3. A higher prevalence is associated with darker colours, while the lightest yellow indicates unsatisfied assumptions. No colour indicates that the area is not under study.

Method-4

During this season, Method-4 could estimate prevalence for only eight of the states examined in the study (see Figure 5.11). This was since these were the states that had sample sizes higher than 30 and non-negative J values. However, by running the MCMC method for all those states, the three states with lowest sample size had a high level of uncertainty, as demonstrated on the trace plots of Figure A.4. In other words, these states were non-convergence. All states, including the ones with the lowest sample sizes had $\hat{R}$ equal 1 and ESS rates ranged between 72% to 90%. Therefore, only the first five states were convergent. The prevalence rates for Andhra Pradesh (65 percent), Telangana (68 percent), Maharashtra (57 percent), Karnataka (56 percent) and Tamil Nadu (70 percent) were exactly the same as in Method-1.

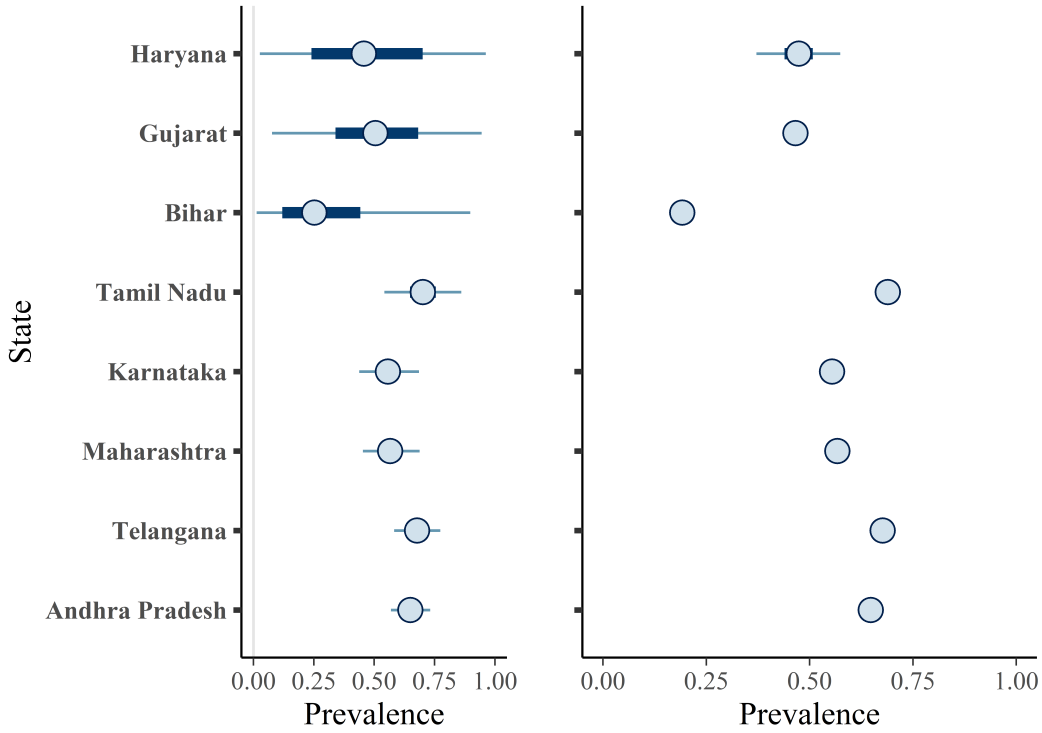


Figure 5.11: Mean posterior prevalence estimates and credible intervals (CrI) of the Indian states during Rabi 2018, where the bold line indicates 50% Crl, and the thin line indicates 95% CrI. The left plot represents forest plot for the original sample size for each state, while the right panel illustrates the forest plot for each state after scaling the sample size by a factor of 10,000.

Increasing the sample size to improve the convergence was only needed for the three states, Bihar, Gajurata and Haryana, as there was an indication that the first five states with highest sample size reached convergence. For consistency, running the MCMC was repeated by multiplying the sample size for each state by a constant factor of 10,000. The right-hand plot of Figure 5.11 demonstrates a decrease in uncertainty across all eight states. Further, all these states successfully estimated by the MCMC method showing convergence. The estimated values remain consistent for the first five states, while for the remaining three states, there is a noticeable improvement, warranting their consideration as reliable estimates. Notably, these values align precisely with those obtained using Method-1 (see Figure 5.12).

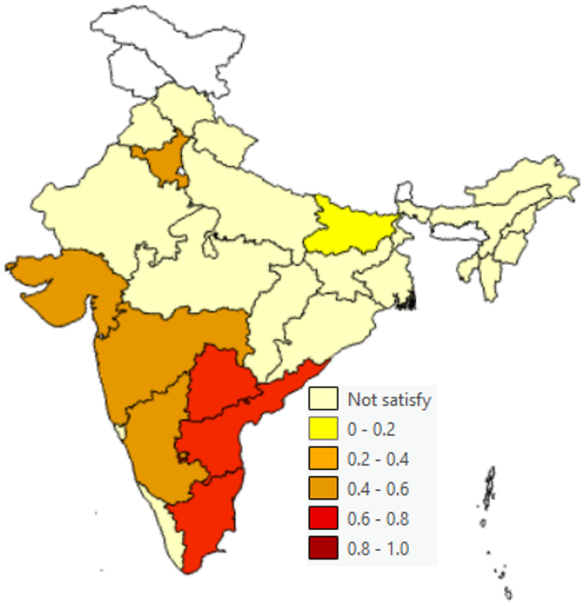


Figure 5.12: Mapping the prevalence of FAW infestation in the maize crop during Rabi 2018 at state level using Method-4. A higher prevalence is associated with darker colours, while the lightest yellow indicates unsatisfied assumptions. No colour indicates that the area is not under study.

5.3.3 Estimation of the prevalence of FAW at state-level during Kharif 2019

Table 5.8: FAW prevalence in India at state level with a total sample size greater than 30 during Kharif 2019, using four different prevalence statistical methods. TP is the number of true positives, and FN is the number of false negatives. P is the total positives, calculated as TP plus false positives (FP). AP is the apparent prevalence. Total refers to the total maize tested in the state. C1, C2, and C3 are the criteria for selecting states. TPr1 to TPr4 are the estimated true prevalence based on Methods 1 to 4. n=10,000 means the sample size for Method-4 is multiplied by 10,000. The (–) indicates that the true prevalence estimations of state are invalid due to not meeting the criteria.

					Method-1		Method-2			Method-3			Method-4	
State	TP	FN	Total	AP	C1	TPr1 (95 CI)	C2	C3	TPr2 (95 CrI)	C2	C3	TPr3 (95 CrI)	TPr4	TPr4, n=10000 (95 CrI)
Maharashtra	2801	1038	5624	0.57	0.50	0.68 (0.67,.69)	0.34	-0.08	-	0.35	-0.06	-	0.68	0.68 (.68,.68)
Karnataka	1292	789	3543	0.45	0.42	0.59 (0.57,.60)	0.22	0.04	0.83 (.67,1)	0.23	0.06	0.8 (.63,1)	0.59	0.59 (.59,.59)
Madhya Pradesh	711	693	2842	0.36	0.30	0.49 (0.48,.51)	0.13	0.13	0.49 (.43,61)	0.14	0.15	0.48 (.42,.63)	0.49	0.49 (.49,.49)
Uttar Pradesh	301	607	2043	0.27	0.11	0.44 (0.42,.47)	0.04	0.22	0.15 (0,.2)	0.05	0.24	0.17 (0,.21)	0.45	0.44 (.44,.45)
Rajasthan	182	450	1601	0.25	0.07	0.39 (0.37,.42)	0.02	0.24	0.07 (0,.14)	0.03	0.26	0.1 (0,.16)	0.43	0.39 (.39,.44)
Telangana	619	286	1457	0.5	0.47	0.62 (0.6,0.65)	0.27	-0.01	-	0.28	0.01	0.97 (.77,1)	0.62	0.62 (.62,.62)
Andhra Pradesh	254	136	648	0.54	0.28	0.60 (0.56,.64)	0.31	-0.05	-	0.32	-0.03	-	0.63	0.60 (.6,.6)
Chhattisgarh	118	122	459	0.38	0.23	0.52 (0.48,.57)	0.15	0.11	0.57 (.49,.76)	0.16	0.13	0.55 (.47,.79)	0.54	0.52 (.52,.52)
Haryana	83	134	459	0.4	-0.04	-	0.17	0.09	0.64 (.55,.93)	0.18	0.11	0.62 (.53,.97)		-
Punjab	62	132	458	0.33	-0.02	-	0.1	0.16	0.38 (.37,.45)	0.11	0.18	0.38 (.36,.46)		-
Tamil Nadu	165	120	449	0.49	0.24	0.63 (0.59,.68)	0.26	0	0.98 (.78,1)	0.27	0.02	0.94 (.74,1)	0.67	0.63 (.63,.64)
Bihar	59	122	423	0.21	0.20	0.43 (0.38,.48)	-0.02	0.28	–	-0.01	0.3	-	0.40	0.43 (.43,.43)
Gujarat	153	84	397	0.5	0.36	0.60 (0.55,.64)	0.27	-0.01	-	0.28	0.01	0.97 (.76,1)	0.62	0.60 (.6,.6)
Jharkhand	66	76	292	0.28	0.35	0.49 (0.43,.54)	0.05	0.21	0.19 (.1,.24)	0.06	0.23	0.2 (.1,.25)	0.46	0.49 (.49,.49)

					Method-1		Method-2			Method-3		Method-4	
State	TP	FN	Total	AP	C1	TPr1 (95 CI)	C2	C3	TPr2 (95 CrI)	C2	C3	TPr3 (95 CrI)	TPr4	TPr4, n=10000 (95 CrI)
Himachal Pradesh	9	40	126	0.17	0.02	0.39 (0.31,.48)	-0.06	0.32	-	-0.05	0.34	-	0.44	0.39 (.36,.4)
West Bengal	21	43	111	0.35	-0.05	-	0.12	0.14	0.45 (.41,.57)	0.13	0.16	0.45 (.41,.58)		-
Odisha	23	13	67	0.46	0.38	0.54 (0.42,.65)	0.23	0.03	0.87 (.71,1)	0.24	0.05	0.83 (.67,1)	0.58	0.54 (.53,.54)
Uttarakhand	12	15	66	0.32	0.21	0.41 (0.30,.53)	0.09	0.17	0.34 (.33,.34)	0.1	0.19	0.34 (.33,.34)	0.46	0.41 (.4,.41)

Method-1

During Kharif 2019 and with applying Method-1 approach, fifteen states met the C0 and C1 criteria. Hence, the FAW prevalence of these states are presented in the map (see Figure 5.13). The map clearly shows that North and South India were comprehensively impacted by the FAW prevalence. Maharashtra had the highest prevalence, at 68 percent. Using this method, the seasonal increase of the FAW prevalence in Karnataka was marginal, yet it increased gradually from season to season. In Kharif 2018, the prevalence rate was 53 percent, subsequently in Rabi 2018, it increased to 55 percent, and during this season it reached 59 percent. This may have been prompted by the great efforts that were made in the state to manage the FAW distribution and preserve maize. Additionally, it is possible that the maize quantities and FAW infestation remained relatively stable in this state across all seasons. In addition, West Bengal, Haryana and Punjab failed to meet C1 criteria during Kharif 2018 and Kharif 2019. One possible explanation could be that there was a more limited amount of maize crop planted during the Kharif season.

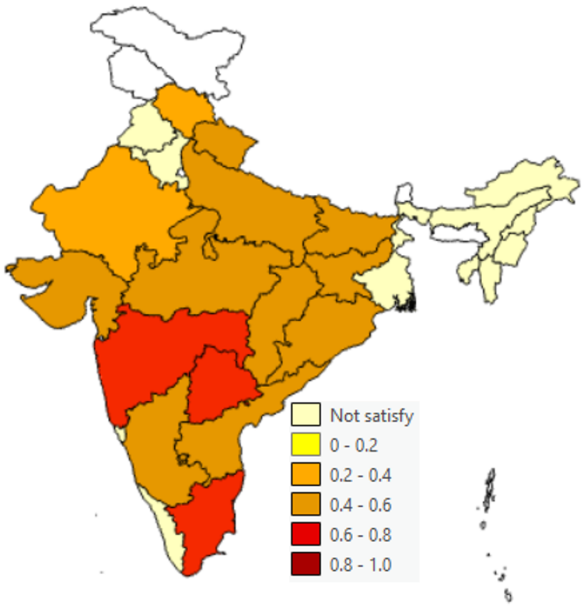


Figure 5.13: Mapping the prevalence of FAW infestation in the maize crop during Kharif 2019 at state level using Method-1. A higher prevalence is associated with darker colours, while the lightest yellow indicates unsatisfied assumptions. No colour indicates that the area is not under study.

Method-2

During the Kharif 2019 season, Method-2 was capable of estimating the prevalence for 12 of the 18 states (see Figure 5.14). Maharashtra, Telangana, Andhra Pradesh and Gujarat failed to satisfy C3 criteria, while Bihar and Himachal Pradesh were not able to meet the C2 assumption. Throughout this season, FAW distribution was exhibited across virtually all northern states, presenting a variety levels of infestation. Further, the prevalence estimates for Madhya Pradesh during this season were approximately identical with Method-1 at 42 percent and Method-2 at 43 percent. In addition, there was a slight difference in the prevalence estimates for Chhattisgarh between Method-1 and Method-2. However, Tamil Nadu exhibited a marked increase reaching 98 percent when compared with Method-1. It is interesting to note that although the sample size and number of positive tests in Karnataka increased over the season, the apparent prevalence remained constant. As a result, the estimated value remained steady, at 83 percent. This may suggest that the level of effective control efforts continued to be stable.

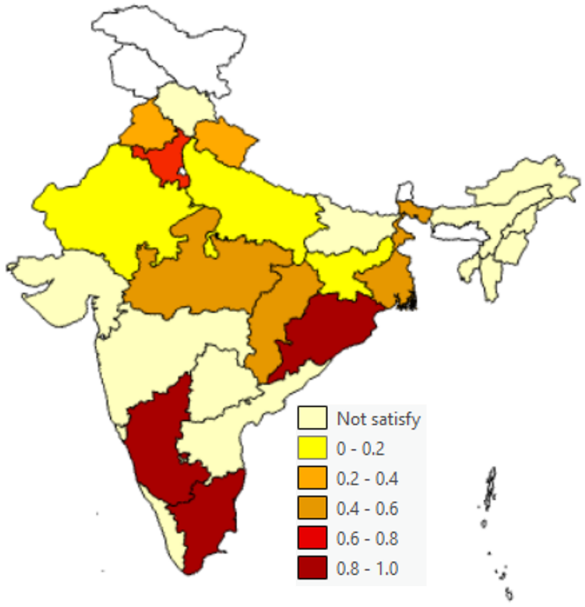


Figure 5.14: Mapping the prevalence of FAW infestation in the maize crop during Kharif 2019 at state level using Method-2. A higher prevalence is associated with darker colours, while the lightest yellow indicates unsatisfied assumptions. No colour indicates that the area is not under study.

Method-3

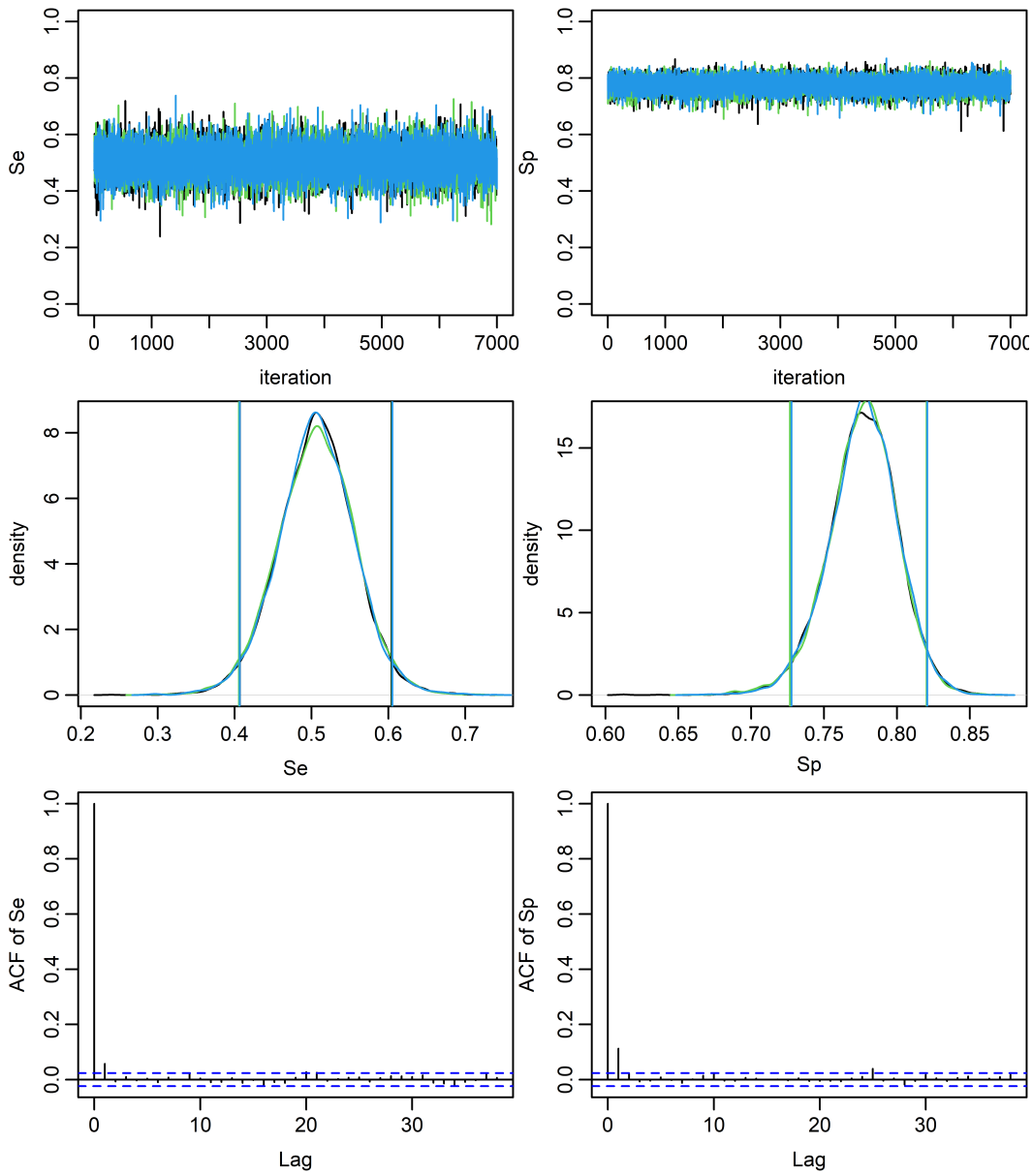


Figure 5.15: Assessing the convergence of three MCMC chains (Post Burn-in) for sensitivity and specificity during Kharif 2019, using trace plots in the first row, density plots in the second row, and ACF Plots in the last row.

Moreover, the MCMC was used to estimate the pooled Se and Sp for Kahrif 2019. Hence, four various diagnostic tools were utilised to test convergence. As shown in Figure 5.15, three of these tools are tthe race plot in the first row, the density plot in the second row, and the ACF plot in the last row. The trace plots were well mixed with overlapping chains, and the density plots show unimodel distributions. Further, the ACF values fluctuate around zero. The $\hat{R} = 1$ and the ESS rate was 100 percent for

Sp and above 57 percent for Se. Therefore, the estimation parameters were a reliable and accurate.

Table 5.9: Point estimates and credible intervals of MCMC summary results for sensitivity (Se) and specificity (Sp) parameters during Kharif 2019.

	Mean	95 CrI
Pooled Se	0.507	0.407, 0.604
Pooled Sp	0.778	0.727, 0.821

Hence, the pooled Se and Sp for states during Kharif season are presented in Table 5.9. By comparing the estimated Se and Sp for this season with the entire period (Method-2), it can be observed that the Se and Sp are almost equal, where the Se for Method-2 is 0.494 and Sp is 0.769. As a result, there was a high chance of obtaining the same estimation prevalence for both methods. However, this method was able to estimate one more state, Telangana (97 percent). Furthermore, Figure 5.16 shows the estimation prevalence based on this method.

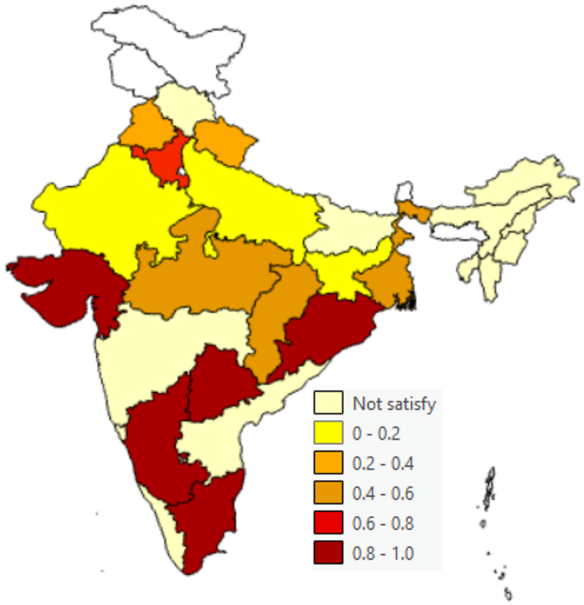


Figure 5.16: Mapping the prevalence of FAW infestation in the maize crop during Kharif 2019 at state level using Method-3. A higher prevalence is associated with darker colours, while the lightest yellow indicates unsatisfied assumptions. No colour indicates that the area is not under study.

Method-4

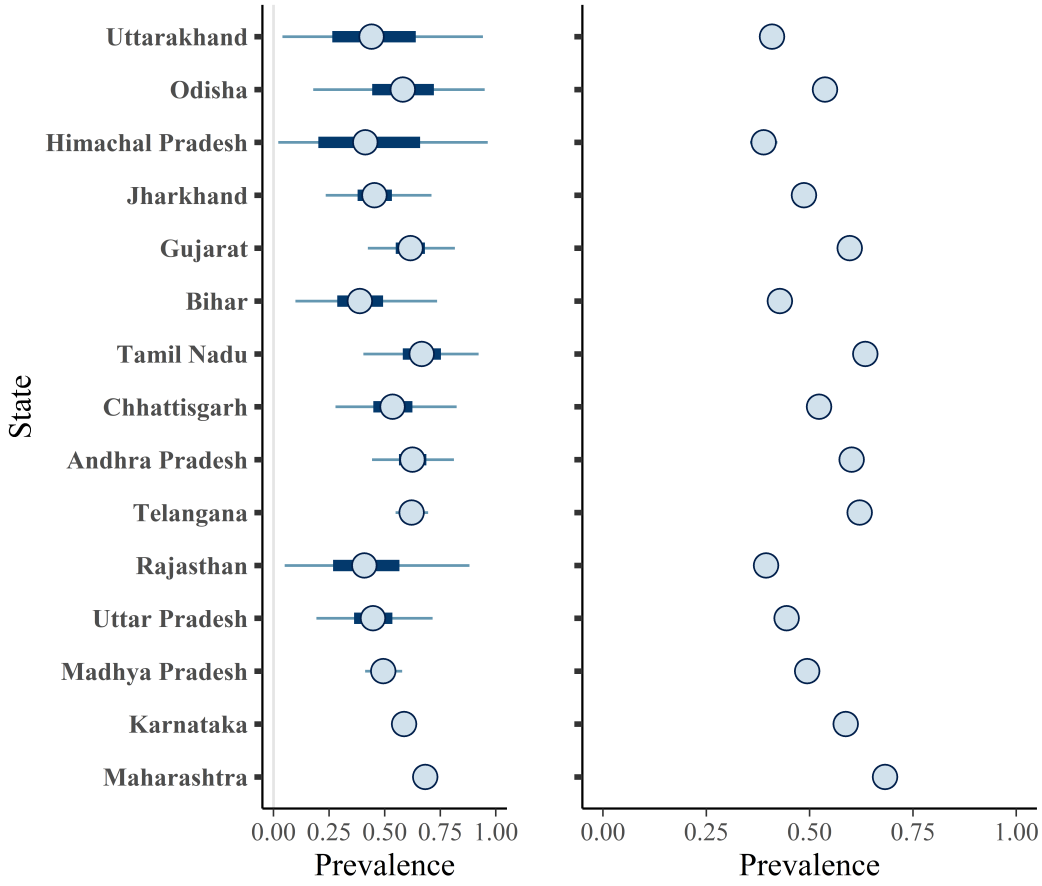


Figure 5.17: Mean posterior prevalence estimates and credible intervals (CrI) of the Indian states during Kharif 2019, where the bold line indicates 50% Crl, and the thin line indicates 95% CrI. The left plot represents forest plot for the original sample size for each state, while the right panel illustrates the forest plot for each state after scaling the sample size by a factor of 10,000.

Furthermore, Method-4 was utilised to estimate the true prevalence of FAW during Kharif 2019 season. Accordingly, 15 states met the criteria C0 and C1. Subsequently, the MCMC process was implemented for those 15 states. Hence, the forest plot shown on the right in Figure 5.17, is a visual representation of the MCMC process, indicating that the method effectively estimated the rates for the majority of states. Specifically, the model successfully estimated the prevalence rates for Maharashtra, Karnataka, Madhya Pradesh, Telangana, Andhra Pradesh, and Gujarat.

Moreover, with less effectiveness and moderate uncertainty, the model also estimated prevalence rates for Uttar Pradesh, Chhattisgarh, Tamil Nadu, and Bihar. However, regarding the remaining states, the model failed to provide estimates. For those with

less effectiveness and moderate accuracy, the model displayed a remarkable degree of similarity with Method-1. Nevertheless, in order to improve the accuracy of the estimations for all states, the MCMC was repeated with a constant factor and a larger sample size 10,000 for each state. Consequently, the updated forest plot shown in the left in Figure 5.17 demonstrates convergence for all states. As a result, the map in Figure 5.18 shows the FAW prevalence based on the repeated analysis. In addition, the FAW prevalence findings were exactly the same as the results from Method-1.

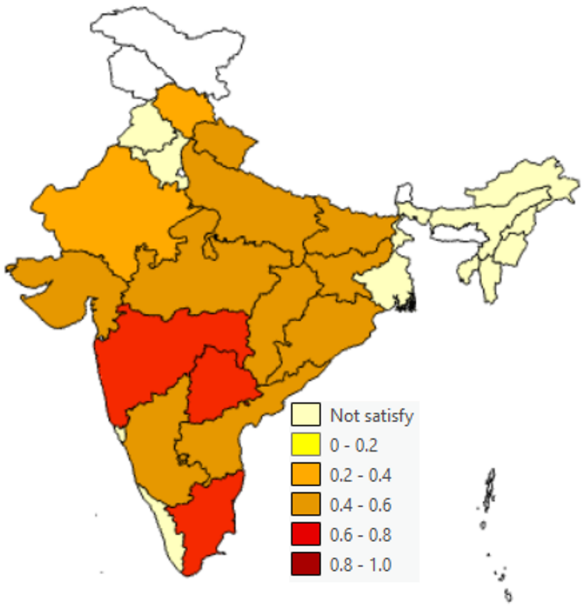


Figure 5.18: Mapping the prevalence of FAW infestation in the maize crop during Kharif 2019 at state level using Method-4. A higher prevalence is associated with darker colours, while the lightest yellow indicates unsatisfied assumptions. No colour indicates that the area is not under study.

5.3.4 Estimation the prevalence of FAW at state-level during Rabi 2019

Table 5.10: FAW prevalence at Indian-state level during Rabi 2019, using four different prevalence statistical methods, where each state has total sample size no less than 30. TP is the number of true positives, and FN is the number of false negatives. AP is the apparent prevalence. Total refers to the total maize tested in the state. C1, C2, and C3 are the criteria for selecting states. TPr1 to TPr4 are the estimated true prevalence based on Methods 1 to 4. n=10,000 means the sample size for Method-4 is multiplied by 10,000. The (–) indicates that the true prevalence estimations of state are invalid due to not meeting the criteria.

					Method-1		Method-2			Method-3			Method-4	
State	TP	FN	Total	AP	C1	TPr1 (95 CI)	C2	C3	TPr2 (95 CrI)	C2	C3	TPr3 (95 CrI)	TPr4	TPr4, n=10,000 (95 CrI)
Maharashtra	751	335	1581	0.56	0.41	0.69 (0.66,.71)	0.33	-0.07	-	0.28	0.13	0.68 (0.66,.76)	0.69	0.69 (.69,.69)
Karnataka	553	202	1093	0.58	0.50	0.69 (0.66,.72)	0.35	-0.09	-	0.3	0.11	0.73 (.68,0.8)	0.69	0.69 (.69,.69)
Andhra Pradesh	460	143	831	0.65	0.40	0.73 (0.69,.75)	0.42	-0.16	-	0.37	0.04	0.9 (0.83,1)	0.75	0.73 (.72,.73)
Bihar	278	111	618	0.5	0.56	0.63 (0.59,.67)	0.27	-0.01	-	0.22	0.19	0.54 (.54,.55)	0.62	0.63 (.63,.63)
Telangana	297	104	525	0.64	0.43	0.76 (0.73,0.8)	0.41	-0.15	-	0.36	0.05	0.87 (0.8,1)	0.78	0.76 (.76,.77)
Tamil Nadu	216	72	405	0.61	0.49	0.71 (0.67,.75)	0.38	-0.12	-	0.33	0.08	0.8 (0.74,.91)	0.73	0.71 (.71,.71)
Madhya Pradesh	180	70	382	0.61	0.33	0.65 (0.61,.70)	0.38	-0.12	-	0.33	0.08	0.8 (0.74,.91)	0.68	0.65 (.65,.66)
Gujarat	128	72	295	0.52	0.37	0.68 (0.62,.73)	0.29	-0.03	-	0.24	0.17	0.59 (0.58,.61)	0.64	0.68 (.68,.68)
West Bengal	125	58	280	0.52	0.45	0.65 (0.6,0.71)	0.29	-0.03	=	0.24	0.17	0.59 (0.59,.62)	0.63	0.65 (.65,.66)
Uttar Pradesh	33	48	133	0.47	-0.17	-	0.24	0.02	0.91 (.74,1)	0.19	0.22	0.47 (0.46,.44)	-	-
Rajasthan	40	35	121	0.48	0.14	0.62 (0.53,0.7)	0.25	0.01	0.95 (.75,1)	0.2	0.21	0.49 (.46,.50)	0.47	0.62 (.61,.62)
Chhattisgarh	25	14	56	0.61	0.11	0.7 (0.57,0.8)	0.38	-0.12	-	0.33	0.08	0.8 (.74,.91)	0.64	0.7 (.68,.7)
Punjab	22	11	52	0.46	0.56	0.63 (0.5,0.75)	0.23	0.03	0.87 (.7,1)	0.18	0.23	0.44 (.4,.47)	0.52	0.63 (.63,.64)
Haryana	16	12	40	0.62	-0.18	-	0.39	-0.13	-	0.34	0.07	0.83 (.78,.97)	-	-
Assam	14	11	38	0.42	0.41	0.66 (0.5,0.79)	0.19	0.07	0.72 (.56,1)	0.14	0.27	0.35 (.26,.39)	0.44	0.66 (.65,.66)

Method-1

By applying this method, and particularly in Rabi 2019 season, it can be noted that the prevalence of FAW varied from 61 percent in Uttar Pradesh to 76 percent Telangana (see Figure 5.19). In Karnataka, there was a gradual increase, which may be linked to the continued attention focused on this state compared to others. In addition, a marginal increase was noted in each state excluding Andhra Pradesh, where a decline from 65 percent to 60 percent was witnessed. Therefore, the FAW infestation surpassed 50 percent in all states, causing concern and a hazardous situation in the agricultural sector, as well as concwening food security, and other related sectors.

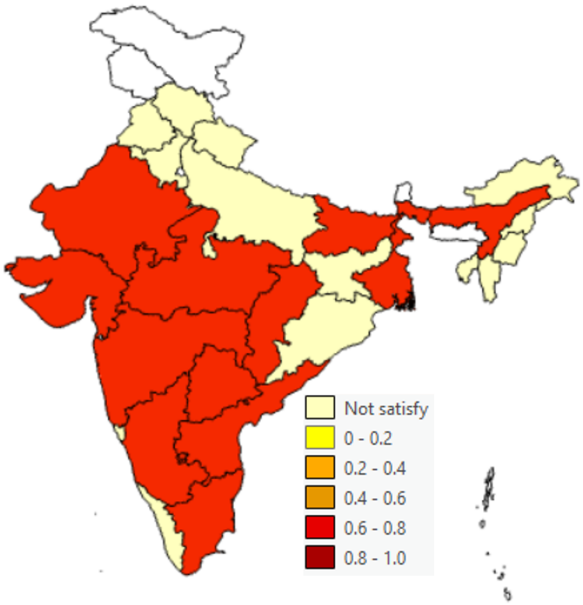


Figure 5.19: Mapping the prevalence of FAW infestation in the maize crop during Rabi 2019 at state level using Method-1. A higher prevalence is associated with darker colours, while the lightest yellow indicates unsatisfied assumptions. No colour indicates that the area is not under study.

Method-2

Most states, during Rabi 2019 and when applying this method, had a FAW prevalence exceeded one. This is because an AP was higher than Se. As a result, the current method was able to estimate the infestation rates of four states; Uttar Pradesh (91 percent), Rajasthan (95 percent), Punjab (87 percent) and Assam (72 percent) (see Figure 5.20). However, these results appeared to be overestimated.

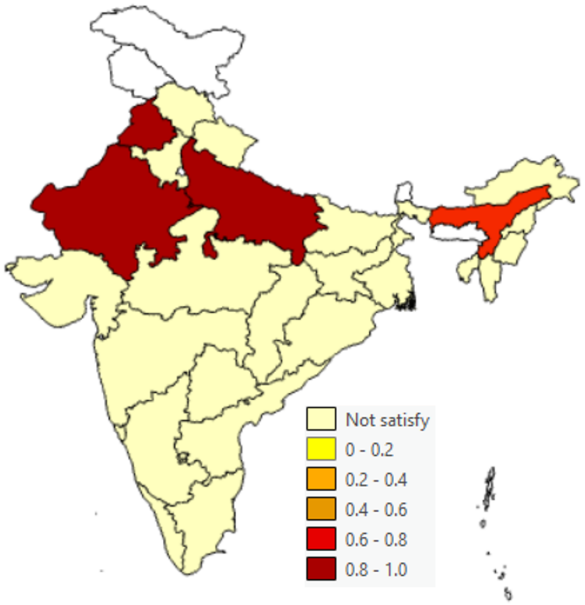


Figure 5.20: Mapping the prevalence of FAW infestation in the maize crop during Rabi 2019 at state level using Method-2. A higher prevalence is associated with darker colours, while the lightest yellow indicates unsatisfied assumptions. No colour indicates that the area is not under study.

Method-3

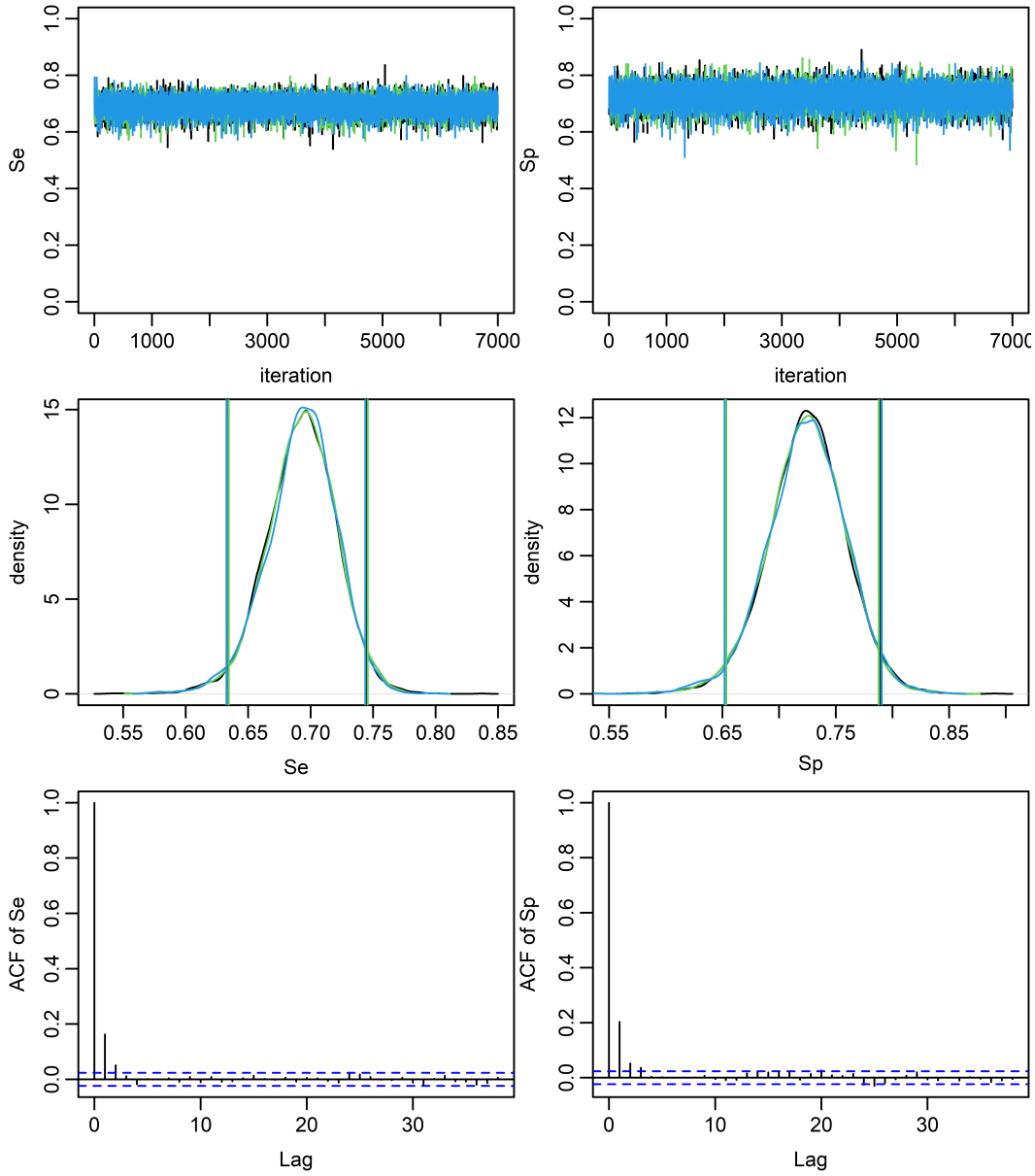


Figure 5.21: Assessing the convergence of three MCMC chains (post burn-in) for sensitivity and specificity during Rabi 2019, using trace plots in the first row, density plots in the second row, and ACF Plots in the last row.

Applying Method-3, the MCMC was implemented to estimate the pooled Se and Sp for Kahrif 2019. Further, to test convergence, four diagnostic measures were applied. The trace plots in Figure (5.21) show well-mixed with overlapping chains. In addition, the density plots within the same Figure show unimodel distributions. Furthermore, the ACF plots at the bottom reveal that the values fluctuated around zero. Moreover, the $\hat{R} = 1$ and the ESS rate was 100% for Sp, and above 57% for Se. Therefore, the

estimation parameters with the MCMC method are reliable and accurate. Hence, the summary of the pooled Se and Sp for states during Rabi 2019 season appears in Table 5.11 which as follows,

Table 5.11: Point estimates and credible intervals of MCMC summary results for sensitivity (Se) and specificity (Sp) parameters during Rabi 2019.

	Mean	95 CrI
Pooled Se	0.693	0.634, 0.745
Pooled Sp	0.724	0.654, 0.790

Remarkably, Rabi 2019 revealed the highest sensitivity and the lowest specificity compared to all earlier seasons. This could indicate that during the Rabi 2019 season, there was a noticeable rise in FAW outbreaks. In addition, farmers may have gained more expertise in using the Plantix app. Further, they were increasingly aware of the FAW dangers. Therefore, they used the app to confirm the presence of FAW. This may have been because the FAW infestation became easier for farmers to recognise. Furthermore, they were motivated to use the app because the app suggests treatments and offers helpful advice to control the FAW. Conversely, the lowest specificity may suggest that farmers increased their focus to take images of maize crops with a high likelihood of FAW infestation.

Accordingly, TPr3 in Table 5.10 and the map provided in Figure 5.22 demonstrate that applying Method-3 during the Rabi 2019 season enabled the estimation of all studied states, with the exception of the states that failed to meet the C0 criteria. By comparison the FAW prevalence findings of Method-2 with this method and during the same seaseon, indicate obvious differences. To start with, Method-2 was able to estimate the prevalence values for no more than four states. Further, these estimations seemed to be exaggerated and unrealistic. This seemed to be due to the values of the C3 criteria in these states were around zero. Although zero is assumed to be acceptable in the C3 criteria, it is possible that a value far different from zero may yield a more accurate estimation. Hence, Method-3 was preferable to Method-2.

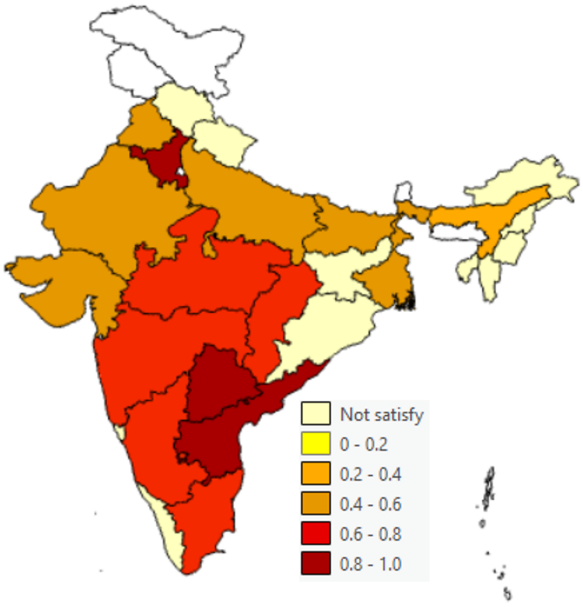


Figure 5.22: Mapping the prevalence of FAW Infestation in the maize crop during Rabi 2019 at state level using Method-3. A higher prevalence is associated with darker colours, while the lightest yellow indicates unsatisfied assumptions. No colour indicates that the area is not under study.

Method-4

Moreover, Method-4 was also applied to simulate the true prevalence of FAW during Rabi 2019 season. Hence, the MCMC process was implemented in 13 of 15 states, which fulfilled the C0 and C1 criteria. Thus, the forest plot on the left in Figure 5.23 shows the MCMC results and the MCMC method successfully estimated for all states except the top four on the plot. It is important to highlight that those four states had the smallest sample sizes.

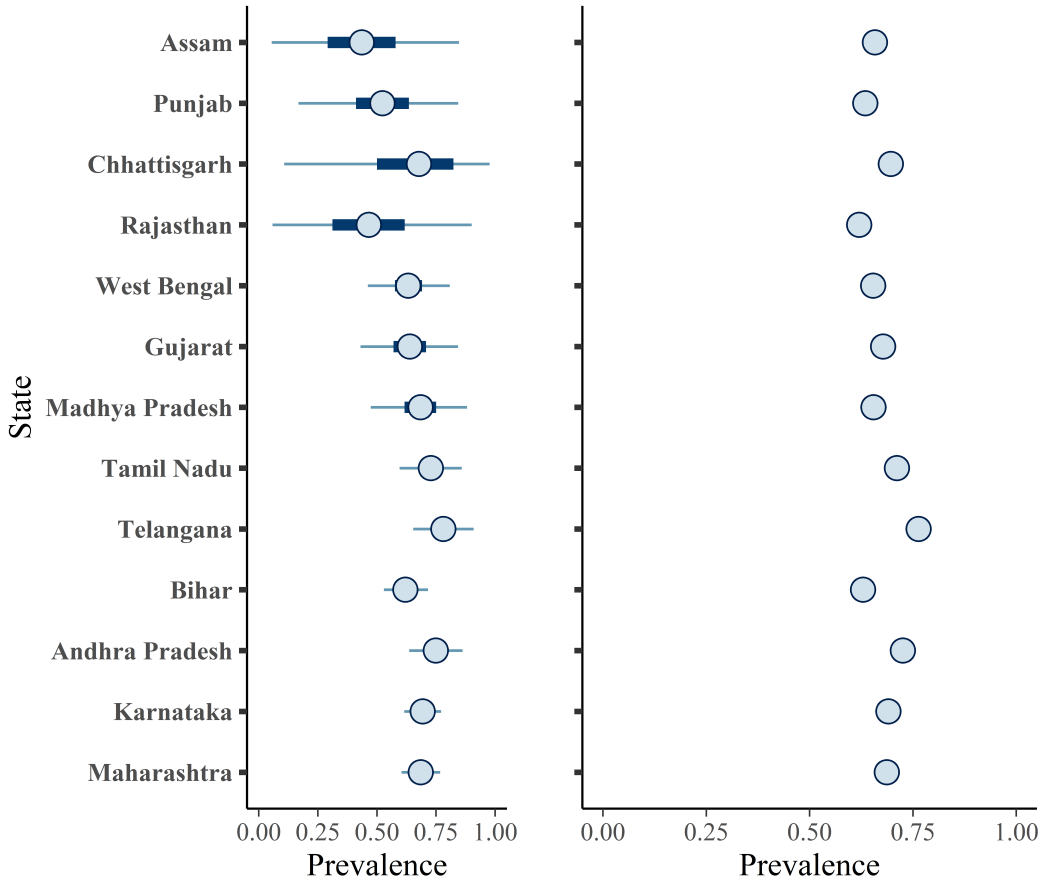


Figure 5.23: Mean posterior prevalence estimates and credible intervals (CrI) of the Indian states during Rabi 2019, where the bold line indicates 50% Crl, and the thin line indicates 95% CrI. The left plot represents forest plot for the original sample size for each state, while the right panel illustrates the forest plot for each state after scaling the sample size by a factor of 10,000.

Therefore, the MCMC was rerun with an increased sample size, by a constant factor of 10,000 for each state. This step was conducted to upgrade the FAW prevalence estimation for all studied states. Consequently, the forest plot, shown on the right side in Figure 5.23, demonstrates convergence for all states. Interestingly, this step of Method-4 and Method-1 yielded the same results. Hence, the map in Figure 5.24 is identical to map in Figure 5.19.

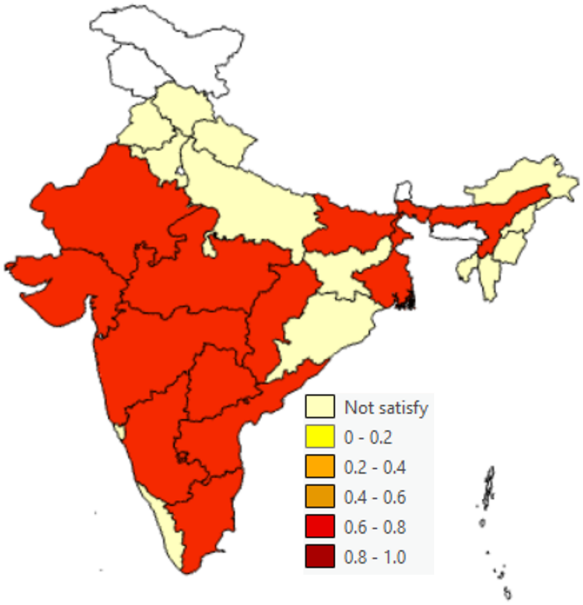


Figure 5.24: Mapping the prevalence of FAW infestation in the maize crop during Rabi 2019 at state level using Method-4. A higher prevalence is associated with darker colours, while the lightest yellow indicates unsatisfied assumptions. No colour indicates that the area is not under study.

5.4 Discussion and conclusion

In this chapter, the primary objective was to estimate the spatio-temporal true prevalence of FAW in maize crop where the spatial scale was state-level in India and the time term was the maize season. Because of the absence of a gold standard and with only a single imperfect diagnostic test, the Plantix app, four distinct statistical methods were applied to identify the most accurate estimate of the true prevalence. The identification of the most accurate method is here determined by the highest method consistency and capacity to produce satisfactory outcomes. The choice is also influenced by a comparison of the estimation values in this study with results from our prior findings or/and prior research studies.

Both Method-1 and Method-4 yielded identical estimation true prevalence values over time and space, when the sample size of Method-4 was augmented to 10,000. Thus, when referencing Method-1/4, it indicates both Method-1 and Method-4. On the other hand, Method-2 and Method-3 presented divergent results from each other. They also exhibited different values compared to Method-1/4. To sum up, while Method-1/4 had the same value, Method-1/4, Method-2, and Method-3 differed. In the following

lines, we compared between the methods to find the optimal one.

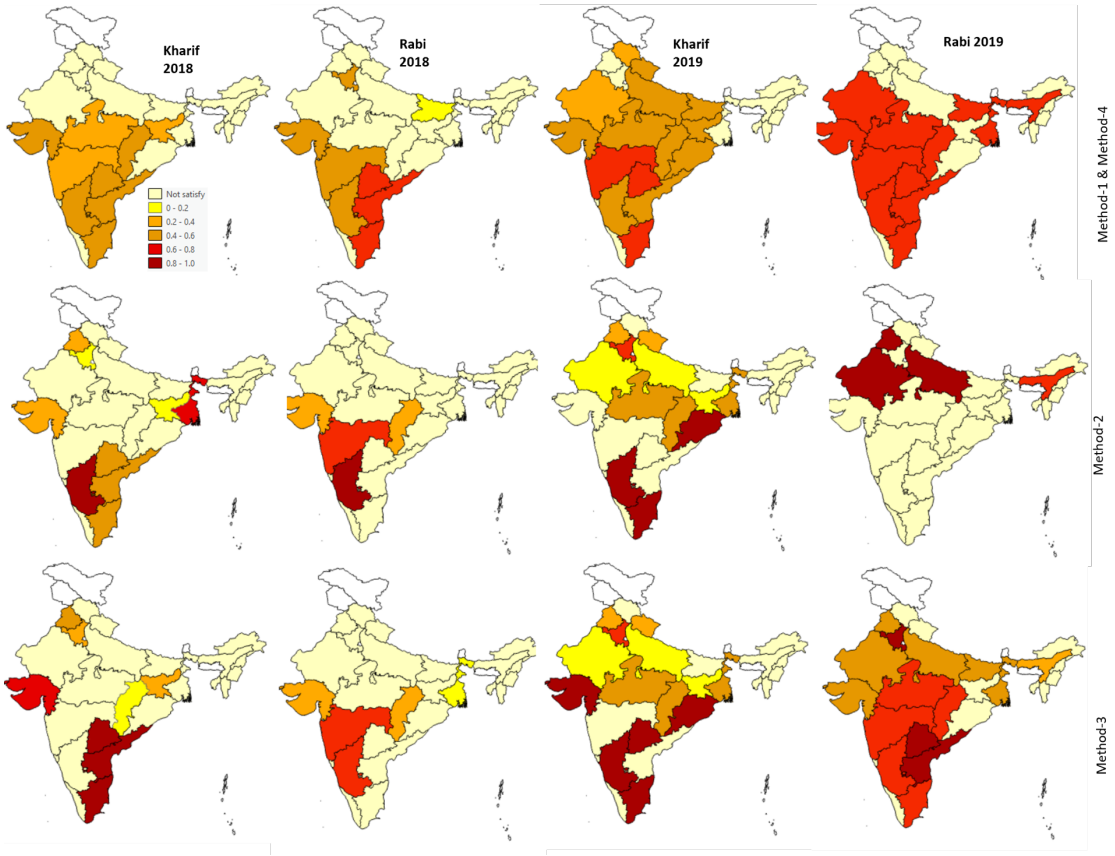


Figure 5.25: Maps of all previous figures to reflect the prevalence of Fall Armyworm (FAW) infestation in maize crops during the 2018-2019 seasons at the state level using Method-1/4. Method-2 and Method-3 in order to compare between them. A higher prevalence is associated with darker colours, while the lightest yellow indicates unsatisfied assumptions. No colour indicates that the area is not under study.

The estimation of the true prevalence in Kharif 2018 in Method-1/4 ranged between 34 percent in Madhya Pradesh and 54 percent in Andhra Pradesh. In Method-2, it was between zero percent in Haryana and 83 percent in Karnataka. In Method-3, there were three states with true prevalence estimation over 80 percent which were Andhra Pradesh,Tamil Nadu and Telangana. Therefore, it is highly unlikely at the onset of the FAW epidemic, that the prevalence surpassed 80 percent. As a result, these results appear to be an overestimation. In addition, during Rabi 2019, Method-2 exhibited to estimate the minimal number of states in comparison to other methods. Therefore, Method-2 highlights possible limitations in accurately representing real-world situations to capture the temporal and geographical variations. Hence, it can be concluded that Method-2 was the least consistent approach and failed to yield satisfactory outcomes.

Table 5.12: Sensitivity and specificity estimation values over seasons.

	Kharif	Rabi
2018	Se=0.376, Sp=0.816	Se=0.543, Sp=0.828
2019	Se=0.507, Sp=0.778	Se=0.693, Sp=0.724

Method-3 demonstrated that the Se and Sp varied throughout seasons (see Table 5.12). This leads to the accuracy of the diagnostic test that can influence the prevalence of FAW infestations across different seasons. In addition, it showed that the Se during Rabi season was better than Kharif in both years. This possibly indicates that the FAW caused more damage in maize during Rabi seasons. Heavy rains during the Kharif season, compared to the Rabi season, impact the availability of Fall Armyworm (FAW) in fields where it cannot be tolerated [21]. Further, Rabi 2019 had the highest Se, almost 0.7, and the lowest Sp. Moreover, the first season, Kharif 2018 recorded the lowest Se, while this value became higher than 0.5 in the subsequent seasons. This can reflect that the FAW prevalence incresed over the time, farmer abilities to recognise FAW increased and their skills in using the app also improved. These findings alignd with our findings of *odds* of FAW in Chapter 3 (see Figure 5.26).

In addition, during Rabi 2019, both Method-1/4 and Method-3 provided meaningful and convincing estimation of the FAW prevalence. They reflected that the FAW prevalence during Rabi 2019 was prevalent across nearly all states, although Method-3 identified that prevalence in the South was more than in the North. On the other hand, Method-1/4 revealed that the level of prevalence in the all states were nearly uniform and ranged from 0.6 to less than 0.8 (red legend). Referring back to our findings about the intensity of the FAW in the Chapter 3 (see Figure 5.26) and comparing them with Rabi 2019 maps in these two methods, it can be seen that in both North and South India the level of intensity was the highest and on average ranged between 2 and 3 odds values, or, alternatively, between 60 and 70 percent. This was consistent with Method-1/4 and not Method-3. This is because Method-1/4 exhibited consistent patterns in both intensity and prevalence, whereas Method-3 did not reflect such uniform pattern in prevalence.

To sum up,

- Overestimation prevalence (Kharif 2018): During Kharif 2018, Method-2 and Method-3 produced prevalence estimates that overestimated the Fall Armyworm (FAW) by over 80%, raising concerns about its reliability.

- Underestimation prevalence (Rabi 2019): In Rabi 2019, Method-2 underestimated the number of affected states compared to other methods (Method-1/4 and Method-3), indicating its limitations in accurately capturing temporal and geographical variations.

- Seasonal Variation Advantage (Method-3): Despite the overestimation, Method-3 had the advantage of demonstrating that the Se and Sp differ over seasons, which could be valuable for understanding seasonal patterns.

- Inconsistency: Given these variations, Method-2 was the least consistent approach and failed to yield satisfactory outcomes across different scenarios.

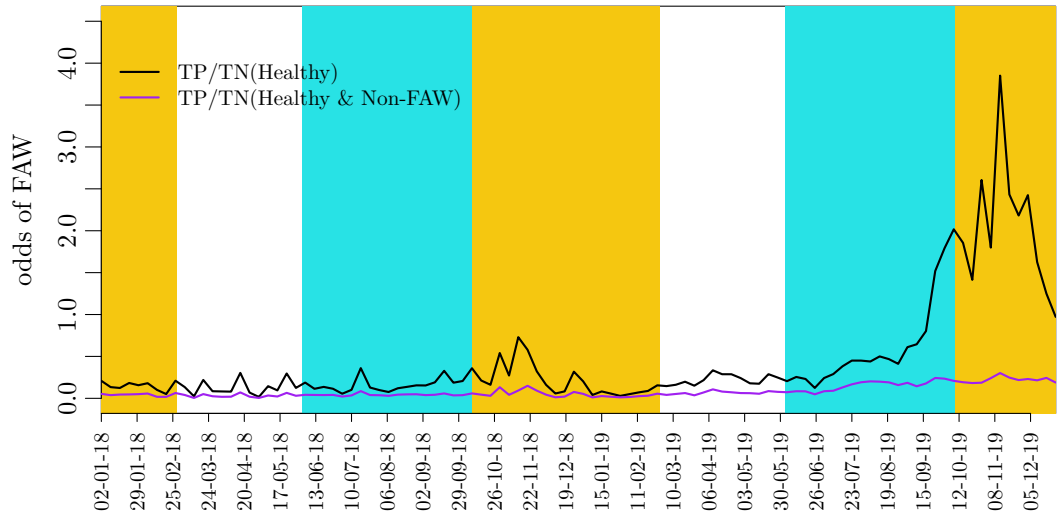
  - Hence,these two methods confirm that directly using the Rogan-Gladen Estimator (RGE) to estimate prevalence by assuming fixed values for sensitivity (Se) and specificity (Sp) can lead to biased estimates of true prevalence.

- Accuracy of Method-1/4: Method-1/4 was the most accurate representation of the prevalence. The method reflects that the FAW concentration was lower in the North area compared to the South during 2018. Moving into 2019, the FAW concentration had increased in both areas. By the end of the 2019, during Rabi, it reached the highest risk level in both areas. The prevalence pattern matches the intensity pattern, and the prevalence map incorporated additional information, including numerical details in state level and across seasons.

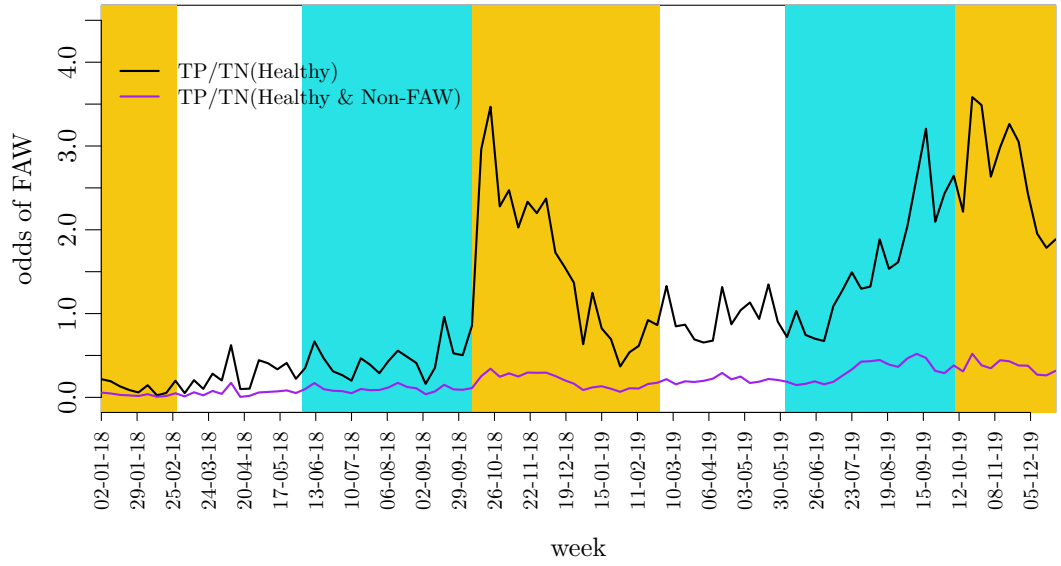
  - Method-1 is the simplest and most direct way to calculate prevalence, making it easy to apply and understand.

  - Method-4 is more advanced and provides useful information. it calculates the pooled Se and Sp values during each season and considers the correlation between these parameters. The Bayesian framework allows to quantify the uncertainty for estimated parameters. Therefoe, the following discussion is based on Method-1/4.

  - However, Method 4 has a limitation that should be highlighted and addressed in future research. The main issue lies in the dependency of the likelihood functions, which requires modification to improve the method's accuracy and applicability



North



South

Figure 5.26: Weekly FAW intensity over the North and South India and over the whole period (2018-2019). These Figures are from Chapter 3.

Moreover, during Kharif 2018, the prevalence map of Method-1/4 shows that the prevalence was in its initial stages. The FAW prevalence was the highest in the South states in Andhra Pradesh at 54 percent, followed closely by Karnataka at 53 percent, Tamil Nadu at 51 percent and Telangana at 50 percent, where it exceeded 50 percent in these states. This evidence demonstrated that the epidemic spread out from these states. Chhattisgarh, Jharkhand and Madhya Pradesh within North India also reported notable FAW prevalence rates. Chhattisgarh had a prevalence of 41 percent, Jharkhand had 36 percent and Madhya Pradesh had at 34 percent. However, these rates may not be

entirely accurate. This is because of their low Youdon index values of 0.01, 0.01 and 0.04, respectively, which round to zero when taken to one decimal place. As a result, these estimates may be closer to random occurrences rather than true representations of the state. It can be concluded that the epidemic originated in the South and subsequently spread to the North. There were reports of FAW outbreaks in the North in Chhattisgarh, Jharkhand and Madhya Pradesh during this season as the initial phase of the epidemic.

It was difficult to compare the prevalence values of the current study with previous studies, due to their limitations to include prevalence analysis. However, existing studies confirmed the presence of FAW in the states that were addressed during this season (Kharif 2018). Their confirmation of FAW occurrence enhance the validity of our findings. During this season, this study revealed the presence of FAW in Gujarat, where 41 percent of maize was infested. Sisodiya et al. [187] documented that the first confirmed sighting of FAW in Gujarat was in September 2018. They observed different characteristics of larvae damage in maize field. These included shot holes on the leaves, consumption the growing cobs and silk, feeding and accumulation of faeces inside the whorls. In my view, these characteristic symptoms of infestation suggests that the FAW may have been present in the field months before September. The Plantix app recorded FAW damage and larvae in Gujarat during January 2018, as shown in the Figure 3.10.

In Maharashtra and during Kharif 2018, this study concluded that the FAW prevalence was 39 percent. Chormule et al. [188] reported that the FAW was identified in August 2018 in 40 days old maize crop and with a recorded infestation level was 20 percent. In addition, this detection can highlight that the FAW was initially observed in maize before sugarcane and sweet-corn. Table 5.13 summarise the first reported month/year of FAW occurrence in the states that show FAW prevalence based on Method-1 and during Kharif 2018:

Table 5.13: First report of fall armyworm occurrence, *Spodoptera frugiperda* in Indian states that show FAW prevalence based on Method-1/4 and during Kharif 2018.

State	Time	References
Karnataka, South	May-June, 2018	[29, 189]
Andhra Pradesh, South	August, 2018	[190]
Chhattisgarh, North	August, 2018	[191]
Gujarat, North	September 2018	[187]
Maharashtra	August 2018	[188]
Tamil Nadu, South	before November 2018	[192]
Telangana, South	2018	[189]
Madhya Pradesh, North	March 2019	[193]
Jharkhand, North	June 2019	[29]

Moving to the following season, the FAW prevalence in the South states was significantly high with no less than 55 percent. Gujarat, on the other hand, had a confirmed true prevalence of 44 percent. Further, Bihar and Haryana reported initial prevalence, where the prevalence rate of Haryana was 47 percent and was 19 percent in Bihar. However, the J value in Haryana was 0.01. Therefore, this small value suggests that Haryana prevalence value appeared to be unreliable and the possibility of an overestimation. According to a study conducted by Kumar et al. [194] the first incidence of FAW was documented in September 2019. This means during Kharif 2019. Nevertheless, their report did not necessarily mean the complete absence of FAW during Rabi 2018, because the current study confirmed some FAW observations. Therefore, it only suggests limitations in the prevalence estimation method or the underlying assumptions.

In the first season of 2019, Kharif, there was still a high prevalence in the South, whereas the North observed the beginning of the prevalence rise. Further, the epidemic extended to all states with the exception of the easternmost states. The prevalence rate exceeding 60 percent was identified in Maharashtra, Telangana and Tamil Nadu which they locate in the South. The prevalence of FAW in Himachal Pradesh was estimated to be at 34 percent, and its Youdon index, was 0.02, which is a non-positive value with rounded one decimal place. This indicates that there was FAW during this season, although the estimated prevalence value may not be accurate and represent the state population. Despite the fact that Sharma [180] reported that the first observation of the FAW was in Kharif 2020, meaning that a year after the findings of this study. In addition, according to Suby et al. [29] Himachal Pradesh was not reported FAW by the study timeframe end, August 2019.

Chapter 6

Estimation the spatio-temporal dynamics of Emerald Ash Borer (EAB, *Agrilus planipennis*) in the USA

6.1 Introduction

Non-native forest insects have caused substantial ecological and economic impacts on forest health [195]. Consequently, forest pests pose a severe threat to the productivity and diversity of native ecosystems (such as changing the light availability and air temperature due to dead or dying trees), industries, and property owners [196, 197]. As global trade and travel have expanded between the United States and other countries, forests have been increasingly invaded by non-native species of insects, where tree mortality is the most critical ecological impact [196, 198]. The emerald ash borer (EAB,*Agrilus planipennis* (Coleoptera: Buprestidae) is one of the most destructive invasive forest pests found in the ash species of North America, causing widespread mortality of ash trees in urban landscapes and natural forests [42, 43, 196, 44, 45, 46]. It has killed hundreds of millions of ash trees in North America and is considered a high-impact pest [196, 57].

EAB is a wood borer beetle that feeds on ash trees (*Fraxinus*) and is also reliant upon them to complete its life cycle [41]. It is indigenous to East Asia (north east China, Japan, Taiwan, Korea, Mongolia and the Russian Far East) and was discovered in the southeastern Michigan, USA in 2002 [42, 43, 44, 45, 46]. However, it is suspected

to have arrived ten years prior with solid wood packaging material from Asia and was only identified when ash trees began to die in massive numbers between 2001 and 2002 [42, 44]. This is because *A. planipennis* is extremely difficult to detect and can remain unnoticed for several years after infestation [44].

While EAB causes minor secondary pest damage to Asian ash species and attacks only dead or severely stressed ash trees [44], all sixteen native Ash species in North America are highly susceptible to EAB mortality [44, 49, 56]. The reason for this is that North American ash species lack resistance mechanisms developed through their evolutionary history with EAB, which Asian ash species have evolved over time [49, 199, 200, 56]. The International Union for Conservation of Nature's Red List of Threatened Species lists six North American ash species as "critically endangered or endangered", with EAB as one of their primary threats [201, 202]. This has resulted in hundreds of millions of dollars being lost by municipalities, property owners, nursery operators, and the forest products industry [57]. Consequently, EAB in North America has received more attention than in its native region [203].

Controlling EAB is challenging for several reasons. From a baseline of 25 US states and 2 Canadian provinces in 2015, EAB has rapidly expanded to 35 US states and 5 Canadian provinces in October 2018. Further, EAB control is complicated by difficulties associated with detection and the time it takes between establishment and response. In addition, monitoring invasion fronts requires extensive resources, such as wide-ranging surveys [41]. Controlling and eradicating EAB can also be expensive, and have adverse effects on the environment. For example, nearly US$1 billion per year was predicted to be spent on treating, removing, and replacing ash trees from 2009 to 2019 [58]. Despite these high costs, successful eradication is not guaranteed [196]. The primary objective should therefore be to find ways to control EAB in an effective and budget-friendly manner.

Mathematical or statistical models can be used to implement to such cost-effective surveillance and control options [121, 103, 41]. Data-driven models can be used to predict the spread of insect threats, as well as to assess potential impacts and improve surveillance and control measures [41, 204]. For example, models are useful for identifying or ranking high risk areas where pests and pathogens can establish themselves [103, 41]. Moreover, early detection that can guide timely control efforts is most likely to be effective and crucial for mitigation [41]. Thus, applying statistical models to pre-

dict and undersatnd EAB is crucial for preventing its spread and protecting ash trees [205].

EAB modelling has been the subject of substantial research, almost exclusively since it was recognised as an invasive pest in North America [203]. The numerous studies of EAB have employed a variety of mathematical and statistical models to understand and predict the insect's spread, either at local or reigonal scale. Related models have meticulously accounted for one or more influential factors in the spread of EAB. These factors encompasse ash tree availability or density, EAB presence-absence [46, 58] or abundance [95], the distance between infested and uninfested areas, as well as human and climate-related influences [96]. Most of these studies are focussed on North America, and primarily spatial scales, ranging from cells to subcounties to counties [96, 43]. The majority of studies consider temporal scale, within year to understand EAB either in North America or Europe [97, 96, 43]. The results of these studies have provided valuable insights into spread of EAB as well as next scientific research.

Moreover, ash tree intensity is critical since EAB is more likely to spread in a high intensity area of ash trees [41, 98, 58]. Further, most studies focus on presence-absence data compared with fewer studies of EAB abundance [95]. Spatial scale is another critical factor differentiating modelling of EAB. In studies focussed on the USA, the spatial scale considered ranged from small scale grid cells to entire counties [96, 43]. Subcounty level analyses emphasised the substantial contribution of human actions and near-neighbor invasions [43]. However, the EAB invasion in the USA gave rise to observation of numerous long jumps of over 100 km which may be more reliable when forecasting EAB at county level, rather than at sub-county level (see e.g. Samuel et al. [96]). Within the spatial scale, the ash intensity can be estimated using a fundamental metrics in forestry basal area of ash. For example Prasad et al. [101] measured the ash intensity in Ohi, a state in the USA, within a spatial scale 270 x 270 meter cells, considering the basal area of ash.

Climate represents another essential class of variable to be considered when understanding and forecasting the potential dynamics of *A. planipennis*. The number of growing degree days (GDD) in a definite area and accumulated between specific points in time, is one of the main climatic factors determining the insect establishment and distribution. A certain number of growing degree days are necessary to complete the pest life cycle. Hence, GDD has a significant impact on survival of insects, and deter-

mining the possible dispersal landscape of EAB adults [97, 99]. For example, low heat availability can limit the suitable area for EAB establishment. According to Orlova-Bienkowskaja [97] EAB adults are unable to colonise an area where its growing degree day accumulation across the year it lower than 700 degree-days.

Additionally, a number of previous efforts have highlighted the importance of long-distance dispersal in predicting the extent of EAB invasion [43, 100, 58, 101], because anthropogenic factors have become a weighty contributor to the spread of EAB. Human practices may facilitate EABs to spread at the local and regional level as long as host plants are present and basic requirements are met. For example, people contribute greatly to EAB spread by moving nursery ash trees, infested logs and firewood [41, 51, 42]. There is also evidence that EAB spread is associated with road networks, motor sports [41, 206], campgrounds, wood product industries, and human population density [98]. Therefore, dispersal models have been widely applied in previous research as valuable tools for understanding the spread of EAB in a local and regional areas.

To account for both natural spread of EAB and human-mediated long-distance dispersal, most studies utilise a dispersal kernel. These kernels are functions that determine the rate of movement of individuals from one location to another and typically indicate that local movements have a higher rates than those where a greater distance is involved [207, 41]. In EAB modeling, the negative exponential dispersal kernel (exponential decay function) was most commonly applied [41]. The negative exponential kernel was used to model EAB spread in North America by Muirhead et al. (2006) [43] and Kovacs et al. (2010) [58] and in Europe by Orlova-Bienkowskaja et al. (2018) [46].

Furthermore, Muirhead et al. (2006) [43] used the negative exponential kernel to model the spread of EAB at subcounty level from 2002 to 2003 in order to predict the dispersal rates of EAB in 2004 to 2005 in Michigan, Ohio, Indiana and Ontario which are in North America. Additionally, Kovacs et al. (2010) [58] used the negative exponential kernel to model the spread of EAB in 25 states in the USA. Correspondingly, they applied the function to predict the annual spread probability of EAB infestation. To achieve this, the model was implemented in the study area that was divided into equal grids. Then, they estimated the EAB probability in a uncolonised grid through considering the distance between grid midpoints. Further, Mercader et al. a year earlier in 2009 [100] had fitted a negative exponential kernel model for predicting EAB larval densities at two different sites in Michigan, USA. In addition to negative exponential

kernel function, Orlova-Bienkowskaja et al. (2018) [46] applied also a Cauchy (fat-tailed) model and normal kernel function to predict the EAB spread. They implemented these dispersal functions in European Russia and neighboring countries, which include countries Ukraine, Belarus, Estonia, Latvia, and Lithuania. They concluded that the Cauchy model is the most appropriate for understanding and predicting the EAB spread, when the pairwise distance between locations is greater than 200 km [46]. Note that, Bienkowskaja et al. [46] did not consider a power law kernel, where the best model, Cauchy, can be closely approximated by a power law kernel.

Gravity models and logistic regression models have also been used to study and predict the EAB dispersal. Gravity models consider both species nature and spatial features (such as road networks, campgrounds, wood product industries, and human population density, infrastructure), and are used to forecast long-distance dispersal [98]. Muirhead et al. [43] applied the gravity model to estimate long-distance dispersal of EAB risk in Michigan, Ohio and Indiana in USA and Ontario in Canada by including three factors. The factors were; (i) the amount of firewood transferred from the epicenter to provincial parks where campers regularly visited. Moreover, (ii) the number of campsites at each provincial park. In addition, (iii) the road distance between the epicenter of the infestation and the provincial parks. Further, Prasad et al. [101] considered the road network distance between EAB areas and campsites, the quantity of campers journeying from EAB locations to Ohio campsites and the traffic density in their gravity model. Furthermore, Muirhead et al. [43] use the logistic regression model, logit model, considering the human populaiton size in non-invaded areas and distance from the epicentre. They conclude that the model significantly improves with inclusion of the human population size.

Although a range of variables and factors were addressed in previous research using models of varying complexity, all of them enhanced the ongoing scientific research, forestry knowledge and governmental information. For example, some models simply considered a distance based kernel function as a factor to predict the EAB spread [46, 58, 43]. Some used only climatic variables in climate-based ecological niche models, such as maximum entropy modelling (Maxent) and genetic algorithm models (GARP) models. They were used in two studies [102, 45] to predict the EAB spread by determining the suitable climate for EAB in North America. On the other hand, Prasad et al. [101] used a complex mechanistic model to estimate and forecast the EAB spread in Ohio, USA.

Their model is a spatially explicit cell-based and a combination of two components: i) a short spread model, which reflected the EAB natural flight, and ii) a long distance model, which simulated the long distance spread due to human practices. The factors considered important in the long distance model are traffic density on major roads, wood products weights, population density and campgrounds.

Recent years have seen a considerable increase in the spatio-temporal data presented in a high resolution. Therefore, this data can be used in statistical and mathematical modelling to facilitate the understanding of individual-to-individual transmission of the invasion process. It also helps to predict the outcome of future colonisations [208, 103, 209]. Spatio-temporal modeling approaches can extract maximal information from data. In addition, they provide an understanding of the underlying dispersal mechanisms, interactions between species, and spatial heterogeneities in the suitability of land for colonisation [210, 103, 209].

The generic colonisation–dispersal model proposed by Catterall et al. [103] is a stochastic spatio-temporal model. It is a generic model designed to be applied when investigating a diverse range of invasive species in a different geographic regions. It aims to estimate and predict the spread of a specie through a space and over time. Although it was orignally applied to invasive vascular plants, the strength of the model is that,

- It accounts for uncertainty in a colonisation time, since the exact time of colonisation is often unknown or uncertain.
- It is fleixible where it can assess multiple hypotheses and include different climatic and envirommental covariates related to the species spread.
- It can forecast spatial heterogeneity in the colonisation time for a study space which is divided into grids, and predict the number of colonised grids over a temporal scale.
- It is an effective resource to develop management strategies in order to limit the dispersal of non-native species.

Additionally, while the previous studies in the USA included temperature as a climate variable in estimating EAB spread at small spatial scales (e.g. within counties), further research is required to replace temperature with GDD in order to predict EAB spread at larger spatial scales.

Therefore, here the model is adapted to study EAB spread in the USA. To our knowledge this study is the first to apply the generic colonisation–dispersal model in EAB study on the whole of the USA. Additionally, while the previous studies in the USA included temperature as a climate variable in estimating EAB spread, further research is required to replace temperature with GDD in order to predict EAB spread at larger spatial scales (e.g. counties). This chapter bridges these gaps by applying the methodology developed by Catterall et al. [103] to study EAB spread in all counties in the USA in a way that accounts for potential environmemtal covariates such as landscape data describing where the ash host is present and its intensity. Climate variations such as GDD, and the EAB distance dispersal, under the effect of natural spread and anthropogenic role. Hence, this study investigates the spatial and temporal spread of invasive species across USA counties.

In the following, we introduce the study area (7.3.1), the data (7.3.1), and the model (7.3.2). We estimate the model parameters to simulate the spatio-temporal dynamics of EAB and investigate the role played by climate and non-climate factors to better assess their relative importance to the invasion dynamics (7.3). We then discuss how our study findings can contribute to a better undersatnding of the geographical distribution of EAB in the USA, useful to plan sustainable management against EAB (7.4).

6.2 Materials and methods

6.2.1 Biology and data

EAB observation data

The United States Department of Agriculture Animal and Plant Health Inspection Service (USDA APHISUS) collects data on the presence of EAB at USA county level. Hence, this provided us with the dataset of the first confirmed observed EAB occurrence, in a month-year temporal scale between 2002 and 2020. The USDA APHIS delivered this data on February 5, 2020, and it matched the data available at the cooperative EAB project on February 3, 2020 (www.emeraldashborer.info). Further, each entry consisted of co-ordinates of the EAB occurrence with county name and along with the observation time. It is worth mentioning that the times provided are not the colonisation times, as these should predate the observation times. However, the data did not propose information about the relationship between the colonisation and observation

processes. Moreover, the USDA APHIS assums that once a county becomes infested, it is considered to be infested indefinitely.

Further, if the latitude-longitude coordinate was missing for a record, the geometric centroid of the county was used instead. In addition, if there was more than one record in a county, only the oldest one (first occurrence) was considered. There are 3085 counties in the USA, while the observations recorded were from 1065 counties. Figure 6.1 shows the first observed occurrences of the EAB in the counties between 2002 and 2020, with each observation corresponding to a specific year. Yet, due to the presence of only one observation for the year 2020, this data point was not included in the further analysis. This could be due to the COVID-19 pandemic. Hence, the probability of detection of occurrences was believed to be far lower than in previous years. Thus one observation is not enough to represent EAB occurrences in that year. Thus, it may result in an underestimated model and consequently, reduced model accuracy. Accordingly, by reviewing the remaining data points, it generally appears that upcoming county colonies of EAB tend to be clustered close to earlier colonies. Further, it can be seen that the EAB colonies originated in southeastern counties and spread westward and northward. However, EAB distributed more rapidly to the west than to the north (see Figure 6.1).

In this study, the spatial scale of data was determined to be at the county level as the county represents the smallest patch in the data, while the time scale was considered to be year. In fact, the county level was also used by Ward et al. (2020) [211] to predict EAB throughout the USA. It is also worth to undescore that the temporal scale of the data is discrete and annual, but the temporal scale of the model is continuous. To sum up, county/annual scale is useful because, at this scale, patterns of spread are readily observable.

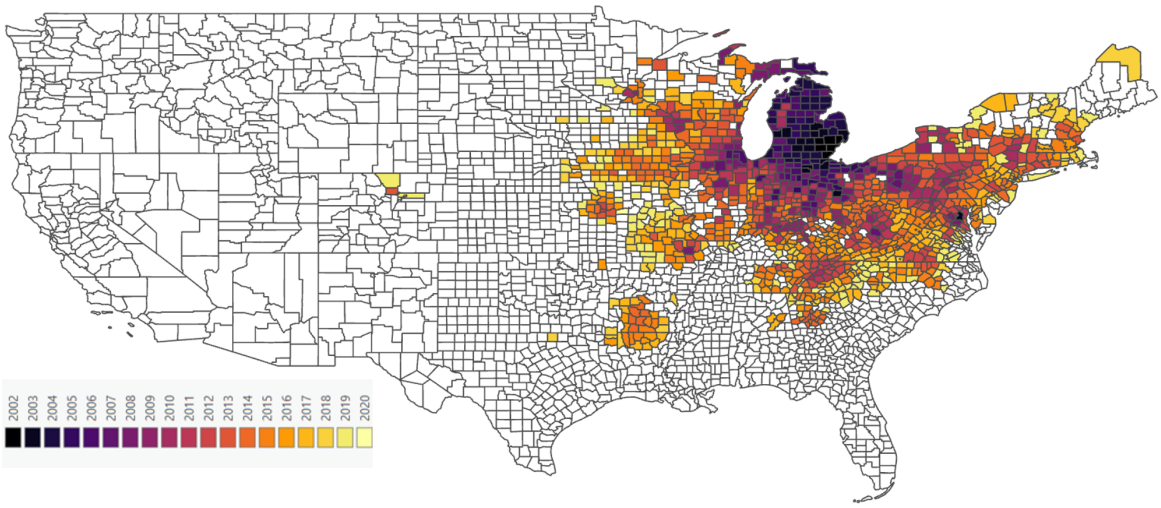


Figure 6.1: First detection of EAB in the USA counties between 2002 and 2020, where the different colours represent colonisation observation time within different years, based on the USDA APHIS data. The map scaled is 1cm=200 km.

Ecoclimate variables

This chapter aims to model the invasion of EAB in terms of appropriate covariates that affect its capability of spreading.

- Environmental covariate: proportion of land in county covered by ash (spatial scale).
- Climatic covariate: Mean for each county (from the beginning of 2002 through the end of 2020) of annual growing degree day

Environmental variable

At the beginning of this research, a data source for ash species intensity in the whole USA was unable to be found. As a result, the deciduous broad leaf trees (DBL) forest was used as a reasonable proxy for ash intensity, which was also be used by Benjamin [212]. However, after a period of time, the ash basal area data was found, which was subsequently used to replace the DBL data. Therefore, in the following lines, the methods to extract both datasets are presented to offer a transparent overview for each of them.

Land-cover by deciduous broad leaf trees (DBL) was considered as a variable to represent ash density. Thus, land cover data were derived from the Copernicus global land service: Land cover 100m: collection 3: epoch 2019: Globe, date of access: 03

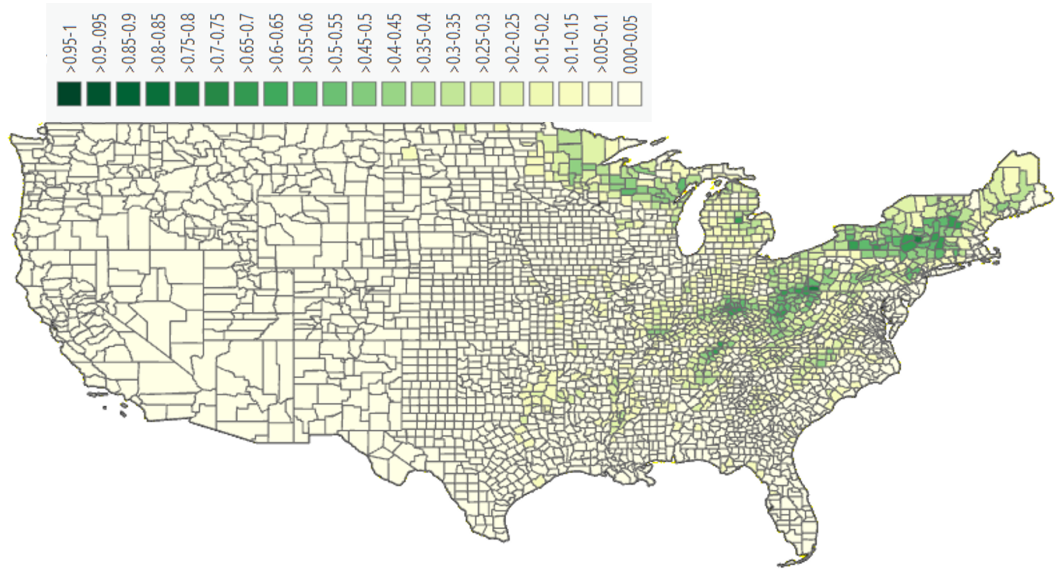
March 2022. This map provides the spatial extent of each forest type (broadleaf forests, coniferous forests, mixed forests, and other types) in each grid square (20° × 20° latitude × longitude degrees). The ArcGIS Pro (version 2.8, ESRI, Redmonds, CA, USA) was used to calculate the DBL proportion area in each county, by summing the DBL area in all 100X100m grid cells within a county and dividing by the county area to end up with DBL a proportion. This covariate does not reflect the exact ash intensity in each county, but hould correlate, perhaps a higher DBL proportion value may be a good indicator to represent a higher ash intensity, and a lower DBL proportion would indicate the opposite.

The second data for ash density in each county was calculated by using ash basal area. Basal area is frequently applied in forestry for calculating tree density and productivity. It is a measure of tree density by calculating the total cross-sectional area of tree breast height (DBH), the diameter of a tree trunk around 4.5 feet (1.3 meters) above the ground, of all live trees within a specific area. The area of each tree can be calculated by using the formula; $\pi\, r^2$,where r is the half value of the diameter. It is represented in square units, such as square meters or square feet, per unit of land area, commonly per hectare or acre [213]. Ash basal area is a good indicator of ash density and a good predictor of the potential for EAB infestation, where it was been used by several studies to predict EAB distribution [214, 95].

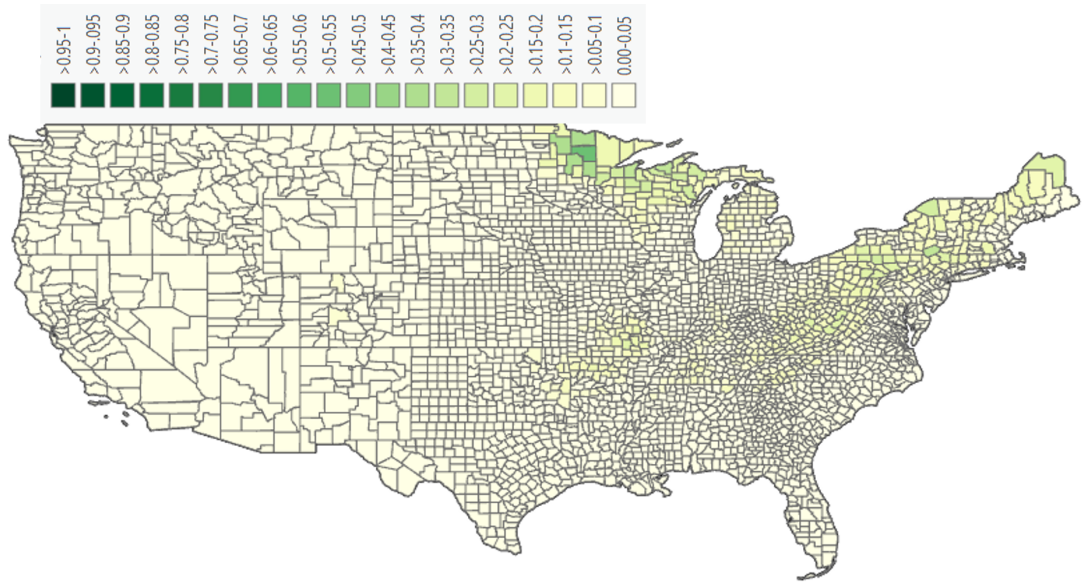
The ash basal area raster map for the contiguous USA was derived from the zipped file 'RDS-2013-0013_RasterMaps_s351-s600.zip' which is available on the USDA Forest Service website [215]. The raster has a resolution of 240 meter-square pixel size grid, each pixel has the value of a live ash basal area. The data represented ash basal area in 2011. The ArcGIS Pro (version 3.1.0, ESRI, Redmonds, CA, USA) was used to calculate the ash proportion area in each county, by summing the ash basal area in all grids within each county, and dividing the figure by county area, to result in the ash probability. Finally, this probability was considered as a model covariate to represent ash density in each county.

It can be noted from Figure (6.2a) and Figure (6.2b) that the USA ash species distribution is not exclusively deciduous broadleaved (DBL) forest. However, the spatial range of ash is wider than the DBL forest. In simple terms, the USA ash species can be found not just in the DBL forest, but also in other areas. Cappaert et al. (2005) [42] reported that the ash species can be located in the forest, urban and suburbs regions.

Moreover, the majority of the ash distribution is located in the east and the south of the USA. Therefore, the selected an environmental covariate to be used for the further analysis, was the ash density map.



(a) Mean deciduous broad leaf trees (DBL) for each USA county by using Copernicus global land service: Land cover 100m: collection 3: epoch 2019: Globe source, where the map was produced by using ArcGIS Pro 3.1.0



(b) Mean ash probability for each USA county by using ash basal area.

Figure 6.2: Comparison of two maps to represent the ash species intensity at the county level in the USA.

Climatic variable

The mean of anual growing degree day (MAGDD) in each county for the period 2002 to 2020 was considered as a covariate in this study, to provide an overview of the

average heat availability in each county. Before calculating the MAGDD, the anuall growing degree day AGDD was calculated from air temperature for each day from the beginning of 2002 through the end of 2020. The air temperature data were obtained from the ERA5-Land Global Atmospheric Reanalysis dataset for each grid square (0.1° × 0.1° latitude × longitude degrees) on the Earth's surface (Copernicus Climate Change Service (C3S): C3S ERA5-Land Reanalysis. Copernicus Climate Change Service, date of access: 17 May 2022). The base temperature (threshold temperature) of 10°C was chosen because it is a standard base temperature previously used in [216, 97]. The $AGDD_{10}$ was carried out using a modified version of the code [97] in the Toolbox Editor in the Coopernicus Climate Change Service (C3S). This result was a NetCDF file, which was opened in R software to convert it to an Excel file and calculated the $AGDD_{10}$ mean for the 18 years in each grid. Finally, ArcGIS Pro 10.8 calculated the mean $AGDD_{10}$ within each county to arrive at the $MAGDD_{10}$. The MAGDD values for each county are shown in Figure 6.3.

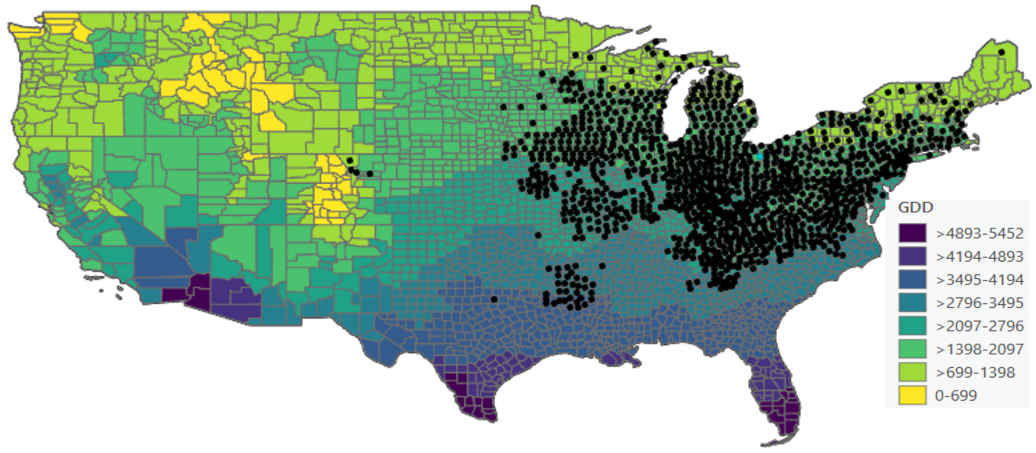


Figure 6.3: Mean of annual growing degree day (MAGDD) from 2002 to 2020 for each USA county.

EAB occurences and ecoclimate factors

The EAB occurences data between 2002 to 2019 in the USA and ecoclimate factors can provide some insights about the most suitable and attractive area for colonisation (see Figure 6.4). Accordingly, Figure 6.4 can suggest a range of possible models for both ash distribution and MAGDD covariates. The y-axis data in Figure 6.4 shows that EAB can inhabit any area with ash trees, regardless of their intensity. Hence, it can be seen that the EAB colonised county with an ash proportion ranging from above zero to the

top value of approximately 0.8.

Turning to the MAGDD covariate, the x-axis data in Figure 6.4 displays that EAB colonised counties had a MAGDD value above 700. Orlova-Bienkowskaja et al. [97] demonstrated that the EAB adult is an unable to colonise an area while its MAGDD smaller than 700. They observed this figure in America, Europe and Asia, where Asia is the native of the EAB. As a result. this value could be the smallest MAGDD that is essential for the EAB establishment of the population. Further, there was a very limited number of EAB observations when the ash proportion was at almost zero and the MAGDD value is seven hundred range. This indicates that it is possible that in some counties the estimated ash figure had been missed or underestimated. Furthermore, there was a small number of EAB cases when the ash proportion was less than 0.3 and MAGDD value was above 3000. Further, it is possible that the optimal and preferred area for EAB survival is when the MAGDD value ranges from 1200 to 3200, although the MAGDD values in across the USA counties range between 0 and 5452.

To sum up, these variations suggest that the MAGDD factor had a significant impact on the behaviour and pattern of the EAB colonisation. Additionally, certain ranges of MAGDD were shown to be more suitable, even with a lower ash density. On the other hand, some MAGDD values were undesirable. Hence, to model suitability as a function of the covariates, the linear equation has potential for modelling the probability of ash intensity factor. In contrast, the normal distribution, gamma distribution and exponential function can be appropriate choices to simulate the MAGDD variable.

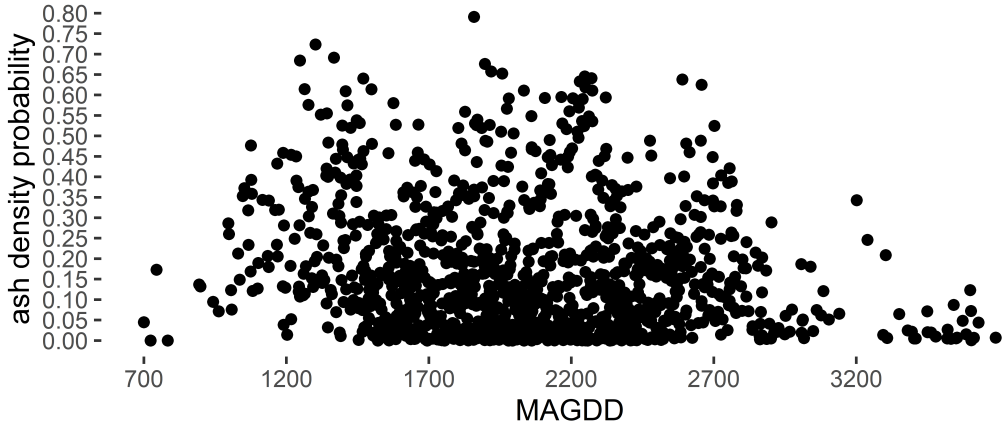


Figure 6.4: The ash probability and mean of annual growing degree day (MAGDD) value for each county colonised by EAB between 2002 and 2019.

6.2.2 Colonisation-dispersal model description

The colonisation-dispersal model is fully specified by equations (6.1)-(6.8) described below. The area to be modelled is divided into discrete, non-overlaping, sites e.g. counties, and the temporal scale is a continuous time (years). Hence the model is a spatio-temporal stochastic process. A stochastic process represents a sequence of random events, where each event is dependent on the previous events and is described by a probability distribution [217]. In the model each county was assumed to be either colonised i or uncolonised j. The colonisation rate r_j of any uncolonised county$j$ at time T is,

$$r_j = \left[\varepsilon + \sum_{i \in C(T)} t_i \, K(d_{ij})\right] s_j \tag{6.1}$$

where t_i is the transmissibility of colonised county i, while s_j is the suitability or susceptibility of site j for colonisation (more details is below). The term in square brackets represents the colonisation pressure from all sources. This colonisation pressure is from both primary/background sources and secondary sources. The primary source is at rate, ϵ, and is outside the area to be modelled such as human density and road network, which may lead EAB to move into a suitable county j [218]. The secondary sources include all counties colonised at time T or earlier, which we represent by the set $C(T)$. The summation is over each colonised site $i \in C(T)$ weighted by both the transmissibility t_i of site i and the dispersal kernel $K(d_{ij})$ which is a function of the distance d_{ij} between colonised site i and uncolonised site j. In the following is the descriptions of the dispersal, suitability and transmissibility functions.

Understanding model components: dispersal functions and their Characteristics

EAB is dispersed locally, or through human-mediated long-distance dispersal within the county or to neighbouring counties. While a mated female EAB can fly up to five kilometers [219], most other EAB adults do not fly more than 100 meters when there are ash trees nearby [58]. A human can, however, disperse EAB over greater distances. Therefore, long distance dispersal of EAB must be considered when studying its spread between counties [43, 58, 220, 46]. The dispersal function is unknown. However, it is clear that the rate of EAB to reach county j from each county i should increase for

counties i that are closer to j and have high transmissibility t_i for the EAB. Therefore, the dispersal function should be a decreasing function as d_{ij} increases. This form of distribution can be illustrated such as a power-law kernel (Equation 6.2) or negative exponential kernel (Equation 6.3), which were used in our analysis. The summation in each denominator is over all possible destination counties j. The mathematical representations of the two kernels are given by;

$$\text{Power-law} \quad K_{pwr}(d_{ij}) = \frac{d_{ij}^{-2\lambda}}{\sum\limits_{g} d_{ig}^{-2\lambda}} \tag{6.2}$$

$$\text{Negative exponential} \quad K_{\exp}(d_{ij}) = \frac{\exp(-\lambda d_{ij})}{\sum\limits_{g} \exp(-\lambda d_{ig})} \tag{6.3}$$

The parameter $\{\lambda,\ \lambda \in (0, \infty)\}$ characterises the decay of the dispersal kernel. It determines how the rate of secondary transmission decreases when the distance between i and j increases. It assumes that the infective challenge decreases as the distance between colonised and uncolonised counties increases. If $\lambda = 0$ in kernel functions (6.2 or 6.3), then any colonised county affects all susceptibles (uncolonised counties) equally, regardless of relative location. Due to this, every susceptible county j is equally likely to be colonised at any given time, leading to a random pattern over time. If λ is large, transmissions are mainly to the nearest neighbours since the infection declines with distance. For any value of λ as the primary source rate goes to infinity dominates secondary interactions, and the spatial distributions of EAB produced by the model appear to be random. Similarly, if primary source is small, and λ is large, transmission is mostly a nearest-neighbor interaction.

The distance metric d_{ij} could be Euclidean or alternative. It is a crucial term in the kernel functions. Orlova-Bienkowskaja et al. [46] calculated d_{ij} using the great circle distances formula, which was also selected for this research. The great circle distance calculates the shortest path between two locations on a sphere, such as the Earth. Therefore, it takes into account the curvature of the Earth. Hence, this calculation provides a reliable and precise calculation comparing with Euclidean distance [221]. The formula can be expressed as

$$d_{ij} = \arccos\Big[sin(lat_i * \frac{\pi}{180}) * sin(lat_j * \frac{\pi}{180}) + cos(lat_i * \frac{\pi}{180}) * cos(lat_j * \frac{\pi}{180}) * cos((lon_i - lon_j) * \frac{\pi}{180}]$$

where lat_i and lon_i are latitude and longitude of county i in a degree unit, and lat_j and lon_j are latitude and longitude of county j in a degree unit. Then, we multiplied d_{ij} by $\frac{180}{\pi}$ to be the final distance d_{ij} in a degree unit. In addition, for computational simplicity, the included colonised counties i in the model were all neighbouring counties $i \in C(t)$ within $d_{ij} \leq 12$ degree. In addition, the model assumed that the colonisation pressure from neighbouring counties i to j is additive i.e. total pressure is a sum of contributions from each county i.

Understanding model components: Suitability and transmissibility

The two sources of covariate information, ash density probability and MAGDD for each county are used to specify a colonisation suitability function s_j, where x represents either i or j county. The s_x is a multiple of both covariate functions, as we assumed that the two covariates were independent. Hence,

$$s_x = h\left(b_x\right) \times f_{\underline{\alpha}}\left(c_x\right) \tag{6.4}$$

Although the ash density probability function $h\left(b_x\right)$ is unknown, it is known that there is a positive correlation between the ash density and the risk of EAB infestation [206]. Therefore, the $h\left(b_x\right)$ should be an increasing function of b_x, where b_x is a normalised ash intensity, representing the fraction of the area that is occupied by ash in a county. As a result, the $h\left(b_x\right)$ can be illustrated as a linear equation, $h\left(b_x\right) = \beta b_x$. Hence, the β is a parameter of ash density probability and measures colonisability (i.e. more suitable for colonisation). Further, it was assumed that it is a constant for all time periods. Moving to trasmissibility t_i, in this work we assumed also that $t_x = s_x$, because both quantities are likely to be related to EAB abundance in the county when become colonised.

In addition, most insects have an upper heat threshold, and exceeding it reduces survival, growth, and reproduction [118]. Therefore, the MAGDD function can possibly be either an decreasing function or a concave-down function. These kinds of functions

can include exponential, gamma, or normal probability density functions. The gamma pdf can adapt to simulate symmetry, left or right skewness. The general formula for all these three functions are as follows:

$$\text{Exponential}: \quad \underline{\alpha} = \alpha_1 \quad f_{\underline{\alpha},Exp}(c_x) = \exp(\alpha c_x) \tag{6.5}$$

$$\text{Normal}: \quad \underline{\alpha} = (\alpha_1, \alpha_2) \quad f_{\underline{\alpha},N}(c_x) = \frac{1}{\alpha_2\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{c_x - \alpha_1}{\alpha_2}\right)^2\right) \tag{6.6}$$

$$\text{Gamma}: \quad \underline{\alpha} = (\alpha_1, \alpha_2) \quad f_{\underline{\alpha},G}(c_x) = \frac{1}{(\alpha_2)^{\alpha_1}\Gamma(\alpha_1)} c_x^{(\alpha_1-1)} \exp\left(-\frac{c_x}{\alpha_2}\right) \tag{6.7}$$

c_x is a scaled MAGDD value (see below) for each county x. A parameter $\underline{\alpha}$ in equation 6.5 can be positive (negative) values, $\underline{\alpha} \in (-\infty, \infty)$, implying warmer (cooler) locations are preferred by the EAB. In the case of the normal and gamma functions, two parameters are involved. In these two functions and during the model-fitting process, one parameter was assigned a fixed value, and the other parameter was estimated by the model. Different potential values and alternation between these two parameters for each iteration of the model were investigated. In the normal, α_1 is the middle of the curve, around the mean, representes the most suitable MAGDD values for the EAB. Subsequently, the tails reflectes the coldest and warmest MAGDD where the EAB impact decreases.

The MAGDD ranges between 0 and high numbers. Hence, it has a much larger and wider range of values than the other covariates. Therefore, scaling MAGDD to values closer to the others could help to minimise numerical instability in the fitting model. The scaled values can be any real number, for both the exponential and normal functions, while for the gamma pdf the value should be non-negative. Six different scaling methods $c_x \in \{c_x^A, c_x^B, c_x^C, c_x^D, c_x^E, c_x^F\}$ (see Table 6.1 for more details) are considered:

$$\begin{array}{ccc} c_x^A = \frac{MAGDD-2344.399}{933.7848} & c_x^B = \frac{MAGDD-700}{5451.229} & c_x^C = \frac{MAGDD}{700} \\ c_x^D = \frac{MAGDD}{1090.246} & c_x^E = \frac{MAGDD}{2100} & c_x^F = \frac{MAGDD}{3100} \end{array} \tag{6.8}$$

Scaling function	Description
c_x^A	It z-scored transformation, where the mean and standard deviation of MAGDD in all the USA counties (3085 counties).
c_x^B	700 is is the minimum MAGDD was observed in the colonised counties, while 5451.299 is the maximum MAGDD was observed in the colonised counties
c_x^C	700 is is the minimum MAGDD was observed in the colonised counties.
c_x^D	1090.246 was obtained by dividing the maximum MAGDD (5451.229) by five. The selection of five was due to the maximum value exceeding 5000
c_x^E	2100 is the middle MAGDD was observed in the colonised counties.
c_x^F	After 3100 there were few observed colonised counties

Table 6.1: Description of different scaling methods.

6.2.3 Comparison of different models for the fitting data

Consequently, there exists a variety of options to model 6.1 in order to simulate EAB distribution. Numerous models were applied in this study. However, Table 6.2 summaries the colonisation-dispersal models that demonstrated the best performance and aid in understanding EAB, which are considered for further analysis.

Model	$f_{\underline{\alpha}}(c_x)$	$K(d_{ij})$
Model 1 ($\exp^A_{\alpha_1} K_{pwr}$)	$f_{\underline{\alpha},Exp}(c_x^A)$	$K_{pwr}(d_{ij})$
Model 5 ($\exp^A_{\alpha_1} K_{\exp}$)	$f_{\underline{\alpha},Exp}(c_x^A)$	$K_{\exp}(d_{ij})$
Model 16 ($N^D_{\alpha_1,.86} K_{\exp}$)	$f_{\underline{\alpha},N}(c_x^D)$ $\underline{\alpha} = (\alpha_1, 0.86)$	
Model 22 ($G^C_{\alpha_1,.5} K_{\exp}$)	$f_{\underline{\alpha},G}(c_x^C)$ $\underline{\alpha} = (\alpha_1, \alpha_2 = 0.5)$	
Model 15 ($G^E_{\alpha_1,.5} K_{\exp}$)	$f_{\underline{\alpha},G}(c_x^E)$ $\underline{\alpha} = (\alpha_1, \alpha_2 = 0.5)$	
Model 17 ($G^D_{\alpha_1,.5} K_{\exp}$)	$f_{\underline{\alpha},G}(c_x^D)$ $\underline{\alpha} = (\alpha_1, \alpha_2 = 0.5)$	
Model 21 ($G^F_{\alpha_1,.5} K_{\exp}$)	$f_{\underline{\alpha},G}(c_x^F)$ $\underline{\alpha} = (\alpha_1, \alpha_2 = 0.5)$	
Model 24 ($G^E_{\alpha_1,1} K_{\exp}$)	$f_{\underline{\alpha},G}(c_x^E)$ $\underline{\alpha} = (\alpha_1, \alpha_2 = 1)$	

Table 6.2: Different colonisation-dispersal models that are considered in understanding the EAB colonisation.

1. To determine the most effective dispersal function to simulate the adult EAB movement, the colonisation model 6.1 is run twice, once with Model 1 ($\exp^A_{\alpha_1} K_{pwr}$) and then again with Model 5 ($\exp^A_{\alpha_1} K_{\exp}$), where the only difference between them is the kernel.

2. The optimal kernel was used in all remaining options for the model 6.1.

3. In the gamma $f_{\underline{\alpha},G}(c_x)$, five models were applied to estimate the shape parameter α_1, while the scale parameter α_2 was fixed. This is because the shape parameter forms the skewness. In addition, the peak (mode=$(shape - 1) \times scale$ when $shape > 1$) of the gamma pdf reflectes the most suitable MAGDD value for the EAB presence. Scaled $c_x^C, c_x^D, c_x^E, c_x^F$ were used for Gamma by dividing the MAGDD covariate by a positive factor to be non-negative.

4. Note that: Other possible values for $\underline{\alpha}$ in both gamma and normal were also examined, but did not include because either did not converge or had lower accuracy.

6.2.4 Bayesian inference

We built our likelihood function by taking advantage of the likelihood equation applied by Cook et al. [121] as well as equation 1, as presented by Catterall et al. [103]. The $f(D \mid \boldsymbol{\theta})$ is the likelihood for complete observation data D, and describes the likelihood of D when given parameter values $\boldsymbol{\theta}$ and the model structure.

$$
\begin{aligned}
f(D \mid \boldsymbol{\theta}) = & \left[\prod_{i \in C(T_N)} \left\{ \varepsilon s_i + s_i \sum_{j \in N_j, T_j < T_i} s_j \, K(d_{ij}) \right\} \right. \\
& \left. \times \exp - \left\{ \varepsilon s_i T_i + s_i \sum_{j \in N_j, T_j < T_i} s_j \, K(d_{ij}) \, (T_i - T_j) \right\} \right] \text{ Colonised term} \\
& \times \prod_{j \in \mathcal{C}^*} \exp - \left\{ \varepsilon s_j T_N + s_j \sum_{i \in N_j, T_j < T_N} s_i \, K(d_{ij}) \, (T_N - T_i) \right\} \text{ Uncolonised term}
\end{aligned}
$$

where $C(T_N)$ is the set of colonised counties i up to T_N where $T_i \leq T_N$, while C^* is the set of uncolonised counties by T_N which are subject to colonisation pressure from all colonised counties at times $T_i < T_N = 2019$, but remain uncolonised. The term εs_j represent background colonisations i.e. those not directly from colonised counties.

The priors $f(\boldsymbol{\theta})$ for different parameters of our model were assumed to be independent, so that $f(\boldsymbol{\theta})$ can be expressed as a product of the priors. The prior distribution for each parameter was assumed to follow a uniform distribution, as no prior knowledge was available, making the uniform distribution an appropriate choice (see Table 6.3). Together these define the posterior distribution and the next step was to estimate the parameters.

Table 6.3: Prior distributions for model parameters.

Model	Parameter	Prior distribution
Model 9 ($N^{B}_{0,\alpha_2} K_{\exp}$)	α	Uniform$(0, 2)$
	β	Uniform$(e^{-100}, 2)$
	λ	Uniform$(e^{-100}, 2)$
	ε	Uniform$(0, 2)$
Model 1($\exp^{A}_{\alpha_1} K_{pwr}$)	α	Uniform$(-3, 3)$
	β	Uniform$(e^{-100}, 10)$
	λ	Uniform$(e^{-100}, 2)$
	ε	Uniform$(0, 10)$
Model 5($\exp^{A}_{\alpha_1} K_{\exp}$)	α	Uniform$(-10, 3)$
	β	Uniform$(e^{-100}, 10)$
	λ	Uniform$(e^{-100}, 5)$
	ε	Uniform$(0, 10)$
Model 16($N^{D}_{\alpha_1,.86} K_{\exp}$)	α	Uniform$(e^{-100}, 5)$
	β	Uniform$(e^{-100}, 10)$
	λ	Uniform$(e^{-100}, 5)$
	ε	Uniform$(0, 10)$
Model 17($G^{D}_{\alpha_{1,.5}} K_{\exp}$)	α	Uniform$(1, 8)$
	β	Uniform$(e^{-100}, 5)$
	λ	Uniform$(e^{-100}, 3)$
	ε	Uniform$(0, 3)$
Model 15($G^{E}_{\alpha_{1,.5}} K_{\exp}$)	α	Uniform$(1, 5)$
	β	Uniform$(e^{-100}, 5)$
	λ	Uniform$(e^{-100}, 4)$
	ε	Uniform$(0, 3)$
Model 21($G^{F}_{\alpha_{1..5}} K_{\exp}$)	α	Uniform$(0, 5)$
	β	Uniform$(e^{-100}, 5)$
	λ	Uniform$(e^{-100}, 4)$
	ε	Uniform$(0, 3)$
Model 22 ($G^{C}_{\alpha_{1,.5}} K_{\exp}$)	α	Uniform$(0, 5)$
	β	Uniform$(e^{-100}, 5)$
	λ	Uniform$(e^{-100}, 4)$
	ε	Uniform$(0, 3)$
Model 24 ($G^{E}_{\alpha_{1},1} K_{\exp}$)	α	Uniform$(0, 5)$
	β	Uniform$(e^{-100}, 5)$
	λ	Uniform$(e^{-100}, 4)$
	ε	Uniform$(0, 3)$

1) Read the observational data. Initialise model parameters with arbitrary values. Initialise each unknown colonisation time in U the set of unobserved colonization times in the time interval $[T_1, T_N]$ uniformly at random on the interval of possible times $[T_{N-1}, T_N]$. Compute initial likelihood.

2) Repeate the following steps sequentially a large number of times, discarding an initial number of iterations as 'burn-in' to remove the effect of the arbitrary choice of

θ_0 and storing all other values of the sampled parameters. We ran three independent chains and in each chain 10,000 iterations following a burn-in period of 2000 iterations. The initial parameters in each chain were randomly drawn based on prior distributions independently. Suppose we were currently at iteration m, then:

1. Evaluate the current likelihood $L^{[m]}$.

2. Generate proposed parameter value λ^* for the parameter λ using the proposal distribution q. The q is to propose the next point to which the random walk might move. It follows a normal distribution $\lambda^* \sim \mathcal{N}\left(\lambda^{(m)}, \sigma^2\right)$ centered on the current value of the parameter, with variances chosen to facilitate convergence of the Markov chain.

3. Recalculate the log-likelihood L^* and prior $P(\lambda^*)$ using the proposed λ^*.

4. Calculated the Metropolis-Hastings ratio, the acceptance probability to a "jump" to λ^* with $\min\left(1, \frac{L^* \times P(\lambda^*) \times q\left(\lambda^*, \lambda^{[m]}\right)}{L^{[m]} \times P\left(\lambda^{[m]}\right) \times q\left(\lambda^{[m]}, \lambda^*\right)}\right) = \min\left(1, \text{Hastings ratio}\right)$.

5. If $u \sim U(0,1) <$ Hastings ratio, then accepte proposed change, so that $\lambda^{[m+1]} = \lambda^*$ and the likelihood is updated to L^*. Otherwise, take $\lambda^{[m+1]} = \lambda^{(m)}$ and the likelihood is unchanged.

6. Repeate steps 1 to 5 for others parameters θ.

7. Repeate steps 1 to 5 for each of the unknown colonisation times T in U, the only difference being that the acceptance probability is $\min\left(1, \frac{L^* \times q\left(T^*, T^{[m]}\right)}{L \times q\left(T^{[m]}, T^*\right)}\right)$, and the new times proposed is also from Gaussian centred the current value.

8. To assess MCMC's convergence, we analysed the parameter trace plots, autocorrelation plots, density plots for both prior and posterior distributions by including all chains, and used the R package coda [129] to calculate the univariate Gelman and Rubin criteria.

6.2.5 Bayesian analysis and interpretation of MCMC samples

The MCMC method generates random samples from the posterior distribution of the parameters, which represent inferences about the parameters and unobserved colonisation times. Statistics such as mean, median, and credible interval (CrI) of a parameter can be estimated from the samples, reflecting the level of belief about its value [121]. A

wider CrI implies increased uncertainty about a parameter's true value. and may also indicate a lack of robustness or significant uncertainty in the data, requiring additional analysis or data. Therefore, decision-makers can use our CrI to determine the uncertainty in the forecast EAB spread and prioritize management actions. Moreover, from a posterior distribution it is possible to obtain a dispersal kernel as a function of distance, or the suitability of a county j to be colonised given MAGDD and DBL. MCMC samples are therefore crucial for decision-making and understanding the robustness of the results.

6.2.6 Evaluating model accuracy with ROC analysis

Assessing model accuracy is essential in real-world applications. It enables assessment of the performance, practicality and reliability of the estimated models. Moreover, it helps with the comparison and definition of the goodness of fit for each model. When assessing the classification models' accuracy, receiver operating characteristic (ROC) analysis is one widely used technique. The ROC curve offers a comprehensive understanding of model practicality through investigating the cut-off of sensitivity and 1-specificity at various classification thresholds (see below for a description of how this was done for the model considered here). Furthermore, the Youden Index (J) (see section 2.4) is frequently applied as a technique to obtain the ideal threshold in the ROC curve. Further, the Area Under the Curve (AUC) is another significant metric determined by the ROC curve. This a single value reflects the overall performance of the fitted model. This integrated approach, through applying the ROC curve, the AUC, and the J values, offers a comprehensive assessment of the model accuracy and practicality.

Subsequently, each colonisation-dispersal model was estimated for the whole study period (2002-2019), and then was used to predict the spread of EAB one year ahead for each successive year from 2003 to 2019. After that, the integrated approach (discuss above) was implemented. The year 2002 was excluded, because the estimated model was trained using data from all previous years to forecast the following year. For example, a fitted model was applied for data in 2002 to predict the colonised counties for 2003. Following that, the fitted model was trained on data from 2002 to 2003 to predict the possible colonistion in the following year, 2004, and so on. This procedure continued until the ending of the time frame, involving 2002 to 2018 to predict the final year in the study period, 2019. To sum up, each part of the process involved including a

one-year step (2002, T-1) to predict the colonistion in the subsequent year (T). Hence, it was possible to identify any temporal trends that impact the model's accuracy. This helped to understand in considerable detail the robustness and performance of model in capturing EAB dynamics.

To build up the ROC curve,

1. The fitted colonisation-dispersal model estimated the colonisation rate (r_j) in year T for every county j that had remained uncolonised in the years preceding the predicting year (T). Subsequently, these estimated rates were compared with observation data at time T, which had a binary outcome with 1 to represent the counties colonised at time T (positive) and denoted 0 for the uncolonised counties (negative).

2. The data points were sorted according to their predicted rates.

3. Then, each threshold d which are typically based on the sorted predicted rates defines the true positive rate (i.e., the proportion of the new county, observed at time T whose colonisation probability exceeded d) and the false positive rate (i.e., the proportion of vacant county, at time T, whose colonisation probability exceeded d) was calculated. In order to do this, the true positive rate can be calculated as the proportion of confirmed colonised counties at time T and their prediction probability at time T exceeding threshold d. On the other hand, a false positive rate is the proportion of ucolonised county higher than threshold d.

4. Next, the ROC curve can be drawn by plotting the true and false positive rates for various d.

5. Then, the AUC calculated as the area under the ROC curve.

6. After that, the single optimal d value was computed by J by subtracting the FPR values from the TPR values obtained at step 3. The optimal d value was the maximum Youden index value.

The above three metrics were calculated using the ROCR package [222] in R [222].

A 2×2 diagnostic table was constructed based on the J value. The TP is the total number of counties when the estimated probability was greater than, or equal to, J and the counties that were truly colonised by EAB at predicting time T. Conversely, the FP

is the number of counties where the estimated probability was at least equal to J, yet the observed counties did not confirm colonisation. In the scenario of TN, the estimated colonistion probability was below the J, and the counties were truly uncolonised by the EAB at predicted time T. In the case of FN, the probability falls below the J, while the counties were colonised at T. Table 6.4 summarises these terminologies:

Table 6.4: Summary of the 2×2 diagnostic table for EAB Scenario at the county-level.

		Observed data	
		Colonised	Non-colonised
Model estimation	Probability$\geq J$	TP	FP
	Probability$< J$	FN	TN

6.3 Results

6.3.1 Visualising model convergence

There were four intrested parameters for each fitted model were estimated using MCMC algorithm which are β, α, λ and ε. The convergence of the three MCMC chains was tested through using visual inspection; trace plots, autocorrelation (ACF) plots and density plots of the posterior and prior distributions for every parameter (Figures: 6.5, in Appendix: B.1, B.2, B.3, B.4, B.5, B.6, B.7), and univariate Gelman and Rubin criteria (or PSRF), (see Table: 6.5). Accordingly, the trace line plots for the three chains (labeled Chain 1, Chain 2, and Chain 3) in all Figures, fluctuate randomly (a random walk) where their are not clear patterns. In addition, the trace plots for each parameter in each chain appear highly consistent around the same values and have an almost perfect overlap. Therefore, the trace plots indicated the convergence and stability for each estimated parameter, and they were likely to represent the true value of the estimated parameter. As a result, each parameter had reached its stationary distribution.

Further, the ACF plots for each parameter in three chains show a highly consistent trend, indicating that the MCMC simulations were stable. Furthermore, the plots show a high degree of correlation in the first few lags and as was expected in a Markov chain, especially at lag 1. However, as the lag size increases, the correlation values gradually decrease for some parameters and significantly decrease for other parameters. Thus, the estimated parameters with a significant decrease in the correlation values indicate that the estimated values are becoming more independent of each other; the MCMC samples

are well-mixed and have converged to the target distribution. Yet, the parameters with a slower rate of decrease in the correlation may indicate that the estimated parameter is more likely dependent on its previous estimates. As a result, additional tools are needed to test the convergence and the reliability of the results.

Moreover, the posterior and prior plots for each parameter, and in each chain, had been plotted as another diagnostic tool for convergence. In addition, the prior and posterior plots visualised the distribution of the parameter before and after running the MCMC simulation, indicating how much information has been obtained from the data. The prior was selected to be uniform for all parameters, assigning equal probability and indicating no adequate knowledge about the parameters. Further, it can be seen that the posterior distributions for each parameter in all three chains are highly consistent and overlap well. Additionally, it can be noted that the width of the posterior distributions are much narrower than the prior distributions. This indicates that the data was more informative and limited the possible range values of the estimated parameters which increased the estimation accuracy and uncertainty in the parameter estimates. It is also a positive indicator and means that the posterior distributions are reliable distributions for the true estimated parameters. To sum up, the prior and posterior distributions for the MCMC simulation converged and the models were a good fit for the data.

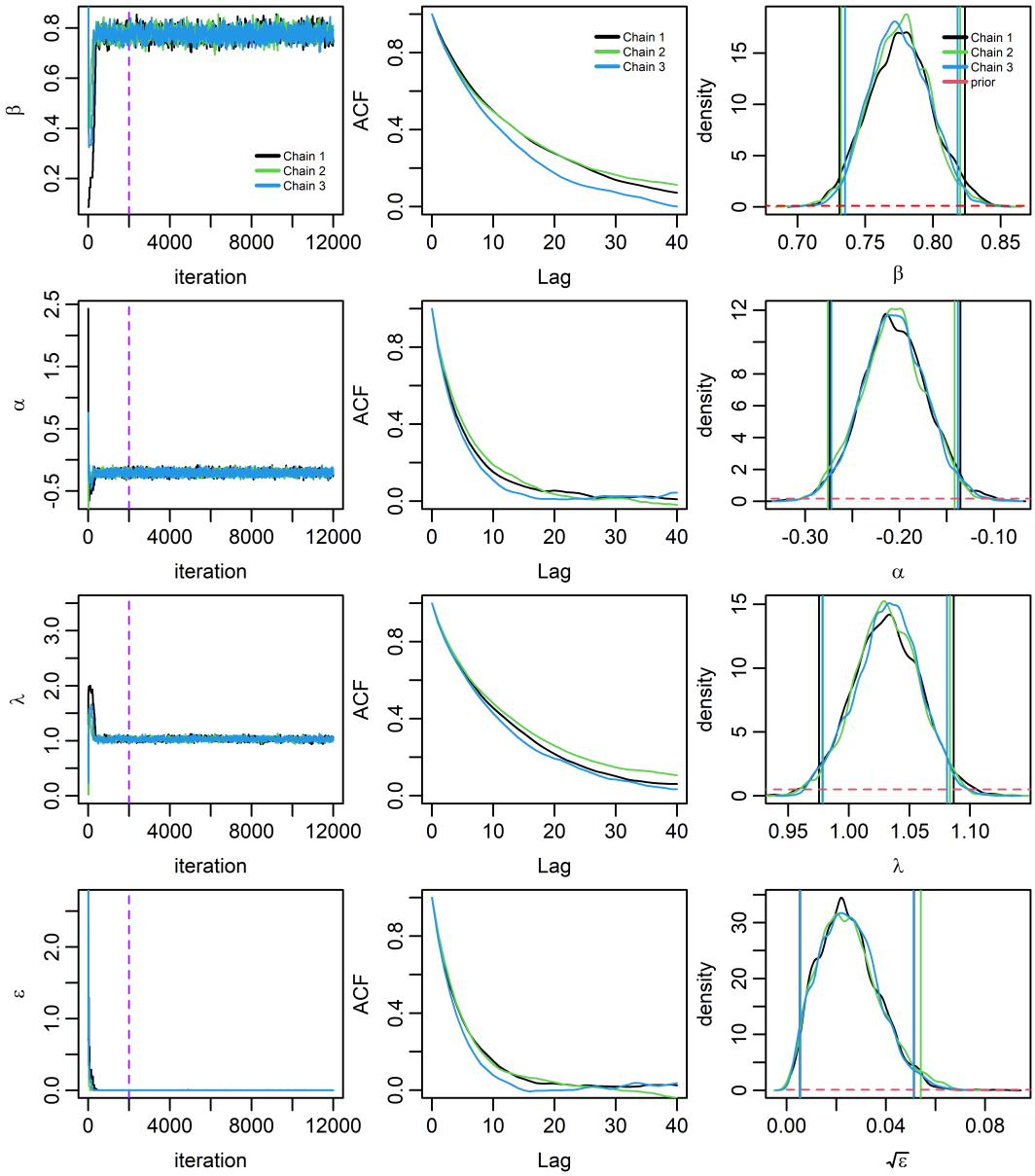


Figure 6.5: Trace plots with a purple dashed line at 2000 (burn-in length), ACF curves, and uniform prior and posterior density plots with vertical lines to represent 95% Credible Intervals (CIs) for three chains of the parameters $(\beta, \alpha, \lambda, \sqrt{\varepsilon})$ of the model 1 $(\exp^{A}_{\alpha_1} K_{pwr})$.

The fourth Bayesian convergence diagnostic tool is the PSRF. The PSRF was calculated based on all three chains of the MCMC. The PSRF point estimate is one or closed to one for each parameter in all models except model 9 $(N^{B}_{0,\alpha_2} K_{\exp})$ which have distinguished values for different parametrs and a value greater than 1.1 for α (Table 6.5). Moreover, the PSRF UCI for all parameters in all models is narrow where there is no value greater than 1.2. Therefore, based on the PSRF point and PSRF UCI, all models converged except, models 9 $(N^{B}_{0,\alpha_2} K_{\exp})$ (see Table 6.5). Overall, the trace,

autocorrelation, and posterior and prior plots all suggest that the MCMC sampler has converged to a good estimate of the parameter except model 9 ($N^B_{0,\alpha_2}K_{\exp}$). Therefore, all models will use for inference analysis except model 9 ($N^B_{0,\alpha_2}K_{\exp}$).

Model	Parameter	PSRF's Point Estimate	PSRF Upper CI
Model 9 ($N^B_{0,\alpha_2}K_{\exp}$)	α	1.04	1.1
	β	1.1	1.11
	λ	1	1.01
	ε	1	1
Model 1($\exp^A_{\alpha_1} K_{pwr}$)	α	1	1
	β	1	1
	λ	1	1
	ε	1	1
Model 5($\exp^A_{\alpha_1} K_{\exp}$)	α	1	1
	β	1	1
	λ	1	1.01
	ε	1	1.01
Model 16($N^D_{\alpha_{1,.86}}K_{\exp}$)	α	1	1.01
	β	1	1.01
	λ	1	1
	ε	1	1
Model 17($G^D_{\alpha_{1,.5}}K_{\exp}$)	α	1	1
	β	1	1
	λ	1	1
	ε	1	1
Model 15($G^E_{\alpha_{1,.5}}K_{\exp}$)	α	1	1
	β	1	1
	λ	1	1
	ε	1	1
Model 21($G^F_{\alpha_{1..5}}K_{\exp}$)	α	1	1.01
	β	1	1.01
	λ	1	1.01
	ε	1	1.0
Model 22 ($G^C_{\alpha_{1,.5}}K_{\exp}$)	α	1	1.0
	β	1	1.01
	λ	1	1.01
	ε	1	1.01
Model 24 ($G^E_{\alpha_{1,1}}K_{\exp}$)	α	1	1.01
	β	1	1.01
	λ	1	1.01
	ε	1	1

Table 6.5: Point estimate and upper confidence interval of the univariate Gelman and Rubin convergence criteria (PSRFs) for each of the model parameters. With values close to one and lower than 1.1 indicate convergence of parameter estimation.

6.3.2 Model estimation parameters

Tables 6.6, 6.7 and 6.8 summarise the posterior medians and 95% credible intervals for the model convergence.

Table 6.6: Posterior medians and 95% credible intervals for parameters.

Parameter	Lower (2.5%)	Median	Upper (97.5)
Model 1 ($\exp^{A}_{\alpha_1} K_{pwr}$)			
α	-0.2741	-0.2062	-0.1384
β	0.7327	0.7753	0.8206
λ	0.9775	1.0316	1.0837
ε	.00002954345	0.0005788704	0.002729169
Model 5 ($\exp^{A}_{\alpha_1} K_{\exp}$)			
α	-0.1575	-0.09121894	-0.02630595
β	0.6803	0.7485	0.8113
λ	1.7054	2.1217	2.5766
ε	0.002866614	0.007088049	0.013562181

Table 6.7: Posterior medians and 95% credible intervals for the estimated parameters with normal function for fitted MAGDD covatiate.

Parameter	Lower (2.5%)	Median	Upper (97.5)
Model 16 ($N^{D}_{\alpha_1,.86} K_{\exp}$)			
α	1.544464	1.611397	1.678556
β	2.464953	2.571904	2.682967
λ	0.4370948	0.5013145	0.5773101
ε	0.00003853288	0.0008878009	0.003674035

Table 6.8: Posterior medians and 95% credible intervals for parameters.

Parameter	Lower (2.5%)	Median	Upper (97.5)
Model 15 ($G^{E}_{\alpha_{1,.5}} K_{\exp}$)			
α	2.3876	2.5390	2.6882
β	1.3511	1.4455	1.5410
λ	1.1683	1.7740	2.3509
ε	0.0024	0.0077	0.01633
Model 17($G^{D}_{\alpha_{1,.5}} K_{\exp}$)			
α	3.8502	3.9973	4.1533
β	2.4687	2.5729	2.6749
λ	0.5292	0.6205	0.7571
ε	0.00007	0.0013	0.0047
Model 21($G^{F}_{\alpha_{1,.5}} K_{\exp}$)			
α	0.3344	0.3646	0.3996
β	0.8972	1.0082	1.1113
λ	1.3441	1.9166	2.4521
ε	0.002982010	0.0083	0.0162
Model 22 ($G^{C}_{\alpha_{1,.5}} K_{\exp}$)			
α	1.4766	1.6115	1.7629
β	3.9791	4.4670	4.9004
λ	1.3002	1.8937	2.5058
ε	0.002771713	0.0084	0.0164
Model 24 ($G^{E}_{\alpha_{1,1}} K_{\exp}$)			
α	0.4929	0.5393	0.5927
β	1.3456	1.4961	1.6531
λ	1.2863	1.8927	2.3879
ε	0.0028	0.0083	0.0160

The parameter values can provide different indication about the EAB colonisation in the USA county level:

- The parameter β is always positive. Hence, it indicates that as the ash densintaty increases in a new county, the likelihood of colonisation by EAB also tends to increase (see β in Table 6.6, Table 6.7 and Table 6.8).

- The parameter α_1 in both ($\exp^{A}_{\alpha_1} K_{pwr}$) model and ($\exp^{A}_{\alpha_1} K_{\exp}$) model is negative. It means that the suitability is a decreasing function of MAGDD. As the heat availabilty in county exceeds 2344.399 (the MAGDD value at the mean of c^{A}_{x} and the intersections point for both models), the suitability of the EAB colonisation becomes less. Conversely, counties with MAGDD values below 2344.399 are more suitable for EAB survival, growth, or reproduction (see α_1 in Table 6.6). Bosed on observed data, Figure 6.4 shows a rare number of colonised counties with MAGDD

value higer than 2344. Further, Model 1 ($\exp^{A}_{\alpha_1} K_{pwr}$) has a larger α_1 value than Model 5 ($\exp^{A}_{\alpha_1} K_{\exp}$), resulting in a steeper decline in suitability as MAGDD increases.

- The parameter α_1 in ($N^{D}_{\alpha_1,0.86} K_{\exp}$) model indicates that the EAB colonisations tend to be highest around 1684 (median) degree-days with potential decreases in suitability as the heat availability falls further from the median. With 95% credibility, the MCMC analysis suggested that the estimated mean of the MAGDD ranged between 1757 and 1830. In addition, the point estimate of the mean ($\mu = 1759$) was slightly higher than the median point estimate (median= 1684). Hence, it indicates that a longer tail extending towards higher MAGDD values. In addition, the right long tail indicated that the EAB was able to survive and spread in warmer climates. Nevertheless, a right-skewed distribution such as gamma could improve the accuracy of simulation of the EAB (see α in Table 6.7).

- The parameter α_1 in ($G_{\alpha_1} K_{\exp}$) models yield distinct estimated values with different scaling methods. For example, the mode of MAGDD in ($G^{D}_{\alpha_1,0.5} K_{\exp}$) is around 1457 ($\alpha_1 \times c_x$), 1554 in ($G^{E}_{\alpha_1,0.5.} K_{\exp}$) and 167 in ($G^{C}_{\alpha_1,0.5} K_{\exp}$). The modes reflect that the EAB colonisation is at the highest and decreases when the MAGDD values are either lower or higher. Models ($G^{F}_{\alpha_1,0.5} K_{\exp}$) and ($G^{E}_{\alpha_1,1} K_{\exp}$) have $\alpha_1 < 1$, hence, they suggest long tail curves and mean that the EAB colonisation rate reduces as the county become warmer (see α in Table 6.8).

- The estimated parameter λ from different kernel functions in models ($\exp^{A}_{\alpha_1} K_{pwr}$) and ($\exp^{A}_{\alpha_1} K_{\exp}$) can compare to understand EAB dispersal behaviour. A lower value in K_{pwr} ($\lambda = 0.98$) suggests that EAB colonisation has higher likelihood at long-distance dispersal. Coversly, the larger value of λ in $K_{\exp}$ ($\lambda = 1.7$) reflects the EAB colonisation prefers short-distance movements (see λ in Table 6.6).

- The estimated parameter λ in all models with ($K_{\exp}$) function are less than 1.34. These values indicate that the dispersal distances from secondary sources of the EAB tend to be short. Further, four models for ($G_{\alpha_1} K_{\exp}$) show different λ values, ranging from 0.52 ($G^{D}_{\alpha_1} K_{\exp}$) to 1.34 ($G^{F}_{\alpha_1} K_{\exp}$). Hence, these values reflect the impact of the scaling methods (see λ in Table 6.8).

To sum up, each model provides unique values for its parameters. Thus, to determine

the optimal model and understand the EAB colonisation, the results of a comparison of various model performances are presented in the following lines.

6.3.3 Models performance

To assess and compare the performance of the models, the ROC curve was employed. In the Appendix, the ROC curves and the corresponding Tables for each year based on fitted colonisation-dispersal model were presented. Several metrics were calculated per year from the ROC curves which were the Area Under the Curve (AUC), the Youden Index (J), the sensitivity (Se) and the specificity (Sp). Hence, the 2×2 classification table was constructed for each year. Under these assessments, model performance improves as false values reduce and the values of all the other metrics increase.

Overview evaluation of annual ROC performance

In general, the yearly AUC exceeded 85 percent in the all models except for the model 22 ($G^C_{\alpha_1} K_{\text{exp}}$) (see Appendix B.2: ROC analysis). The range of the AUC values in the model 22 ($G^C_{\alpha_1} K_{\text{exp}}$) ranged between 99 percent to 75 percent, where the lowest value occured in 2015. Despite this variation, the performance of all models were consistently high. In addition, 2014, 2016 and 2018, recorded the minimum AUC scores in all models, and none of them scored AUC value above 90 in these years. This finding reflected that the root of the lower performance did not relate to the models. Yet, there were unexpected spatial distribution behaviour during those years. In other words, the EAB colonistion in counties in those years might located far away from the exectence once (for example see 2014 in Figure 6.1). Hence, there might be external factors that caused those long range colonisations, such as trasfer of the EAB due to the human activities. Another plausible explanation, experts failed to document the first EAB colonisation counties in those years or in years earlier.

Summary of annual ROC for model performance evaluation

Table 6.9: Summaries the performance of the each model based on the average annual performance .

Model	Average			Cumulative		
	AUC	Se	Sp	FP	FN	FP+FN
model 5($\exp^{A}_{\alpha_1} K_{\exp}$)	0.92	0.95	0.82	7735	79	7814
Model 24 ($G^{E}_{\alpha_1,1} K_{\exp}$)	0.91	0.92	0.824	7918	95	8013
Model 15 ($G^{E}_{\alpha_1,0.5} K_{\exp}$)	0.91	0.93	0.821	8006	77	8083
Model 21($G^{F}_{\alpha_1,0.5} K_{\exp}$)	0.91	0.91	0.81	8057	114	8171
Model 1 ($\exp^{A}_{\alpha_1,0.5} K_{pwr}$)	0.91	0.99	0.79	9008	106	9114
Model 16($N^{D}_{\alpha_1,0.86} K_{\exp}$)	0.91	0.94	0.79	9263	84	9347
Model 17($G^{D}_{\alpha_1,0.5} K_{\exp}$)	0.91	0.94	0.78	9434	80	9514
Model 22($G^{C}_{\alpha_1,0.5} K_{\exp}$)	0.85	0.85	0.75	10556	201	10757

Table 6.9 summaries the overall performance for each model, calculating the average across all years for AUC, Se and Sp, and cumulative number for false terms. Hence, it is obvious that the worst performing model with the highest total false number and the lowest score in the other mertics was the model 22 ($G^{C}_{\alpha_1,0.5} K_{\exp}$). In addition, the specificity values across all models were at all times less than any Se value observed in all models. Consequently, the FP numbers were higher than the FN cases. In other words, the models incorrectly identified substantial numbers of non colonised counties as colonised. Additional relevant covariates integrated in the analysis, might help to improve the performance of the models to understand the EAB colonisation. Furthermore, there may be missing data in this dataset collected by experts. Thus, involving citizen scientists in data collection increase the number of the observations over spatial and temporal scales.

The best model is Model 5 ($\exp^{A}_{\alpha_1} K_{\exp}$) where exhibits the highest average AUC and the lowest falses (see Table 6.9). However, it is worth to mention that all other models show excellent performance, particularly Model 24 ($G^{E}_{\alpha_1,1} K_{\exp}$) and Model 15 ($G^{E}_{\alpha_1,0.5} K_{\exp}$). Therefore, comparison between models provide deeper insights into significant factors influencing EAB colonisation. Hence, the following comparative model evaluations were conducted:

- Comparison 1: Model 5 ($\exp^{A}_{\alpha_1} K_{\exp}$) outperformed Model 1 ($\exp^{A}_{\alpha_1} K_{pwr}$), although they only differ by choice of kernel function $K(d_{ij})$ (see Table 6.9). Hence, this comparison should be on EAB dispersal distance behavior (details in the following sub-section).

- Comparison 2: The four models Model 22 ($G^{C}_{\alpha_1,0.5}K_{\exp}$), Model 15 ($G^{E}_{\alpha_1,0.5}K_{\exp}$), Model 17 ($G^{D}_{\alpha_1,0.5}K_{\exp}$) and Model 21 ($G^{F}_{\alpha_1,0.5}K_{\exp}$) were identical in every aspect except for the scaling method used. Thus, these models can highlight the impact of scaling in model performance (details in the following sub-section).

- Comparison 3: Model 24 ($G^{E}_{\alpha_1,1}K_{\exp}$) and Model 15 ($G^{E}_{\alpha_1,0.5}K_{\exp}$) were identical in every aspect except on their fixed shape parameters, where α_2 was 1 and 0.5, respectively. Moreover, these models are the second and the third best-performing models, respectivly. Hence, they can help to define the MAGDD function, and the relationship between heat availability and EAB colonisation behaviour (details in the following sub-section).

6.3.4 Impact of model components for understanding the EAB colonisation

Comparison 1: Optimal dispersal kernel function

Model 1 ($\exp^{A}_{\alpha_1} K_{pwr}$) with power-law kernel and 5 ($\exp^{A}_{\alpha_1} K_{\exp}$) with negative exponential kernel can be used to compare and visualise the impact of the kernels in model performance. It is well known that the negative exponential kernel function decreases rapidly with increasing distance compared to the power law kernel. This can also be confirmed in this research findings in curve plots in Figure 6.6 where the red and orange curves are of model 1 ($\exp^{A}_{\alpha_1} K_{pwr}$) and 5 ($\exp^{A}_{\alpha_1} K_{\exp}$), respectively. It can be seen that the negative exponential kernel function is a short-range kernel while the power law kernel function is a long-range kernel.

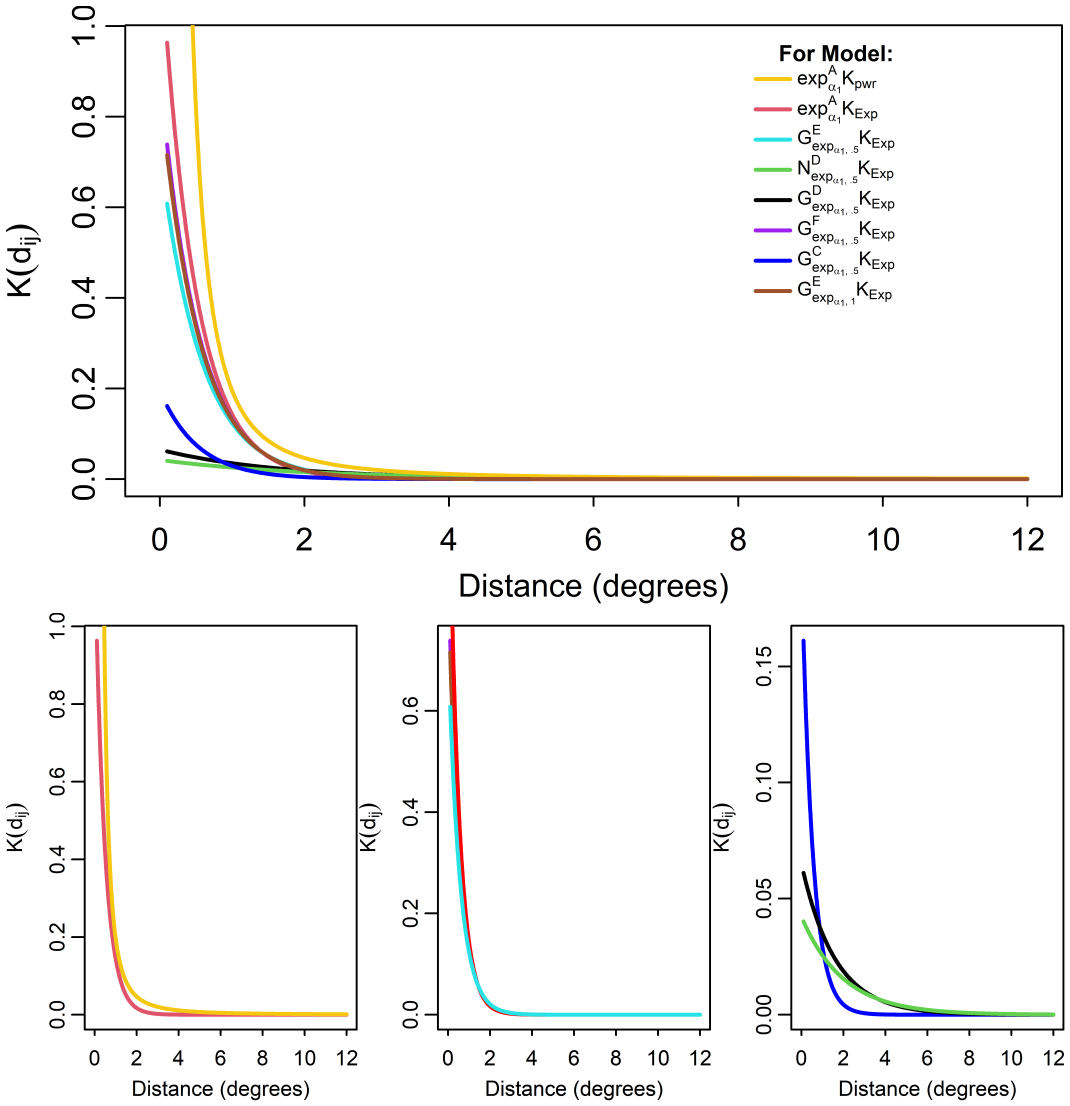


Figure 6.6: Comparison the kernel function distance curves for different models. The Figure at the top includes all kernels. The Figure in the bottom left shows the kernels for Model 1 ($\exp^{A}_{\alpha_1} K_{pwr}$) and model 5 ($\exp^{A}_{\alpha_1} K_{\exp}$) which their models are same except kernel term, allowing for a direct comparison between them. The Figure in the bottom middle displays kernels for models that exhibit performance ranging from excellent to good: Model 5 ($\exp^{A}_{\alpha_1} K_{\exp}$), Model 24 ($G^{E}_{\alpha_1,1} K_{\exp}$), Model 15 ($G^{E}_{\alpha_1,0.5} K_{\exp}$), and Model 21($G^{F}_{\alpha_1,0.5} K_{\exp}$). The Figure in the bottom right illustrates the kernels for the worst-performing models: Model 1 ($\exp^{A}_{\alpha_1} K_{pwr}$), Model 16($N^{D}_{\alpha_1,0.86} K_{\exp}$), Model 17($G^{D}_{\alpha_1,0.5} K_{\exp}$) and Model 22($G^{C}_{\alpha_1,0.5} K_{\exp}$).

This research outcomes found that a noticeable difference in the amount of false cases between these two models, which enhance the role of kernel in understanding EAB dispersal behaviour. The findings confirmed that Model 1 ($\exp^{A}_{\alpha_1} K_{pwr}$) produced higher false rates than Model 5 ($\exp^{A}_{\alpha_1} K_{\exp}$) (see Table 6.9). The model 1 ($\exp^{A}_{\alpha_1} K_{pwr}$) had a total of FP=9008 and FN=106, which was more than the model 5 ($\exp^{A}_{\alpha_1} K_{\exp}$),

which had a total of FP=7735 and FN=79 (see Table 6.9). In other words, the model 1 ($\exp^{A}_{\alpha_1} K_{pwr}$) exhibited a alarming difference of 1300 total number of falses than model 5($\exp^{A}_{\alpha_1} K_{\exp}$). Hence, the negative exponential kernel function ($K_{\exp}$) was a desirable function and improved performance of the understanding EAB distribution.

Both models, for example, were applied separately to the observed data to distinguish the impact of the kernel functions in predicting the presence of EAB in 2014. Yet, it can be applied to any year (Figure 6.7). Hence, it can be seen that the K_{pwr} extended wider than the $K_{\exp}$. Thus, the power law can model the spread of EAB accross wider distances (counties) which may cause an increasing of the number of FP (i.e. predicted non colonised county as colonised). On the other hand, the negative exponential kernel is more responsive to the close counties than far counties. As a result, the $K_{\exp}$ can cause an decreasing of FP. Model 1 ($\exp^{A}_{\alpha_1} K_{pwr}$) produced higher false rates than Model 5 ($\exp^{A}_{\alpha_1} K_{\exp}$), hence, negative exponential kernel $K_{\exp}$ is better than power-law kernel to understand the EAB dispersal distance. To sum up, this analysis highlights the effectiveness of including the kernel function as an a factor to understand the EAB abundance, and dispersal is mostly very localised, but with a small number of very long distance dispersals, something which is best modelled by the negative exponential kernel in combination with a power-low rate of background dispersal.

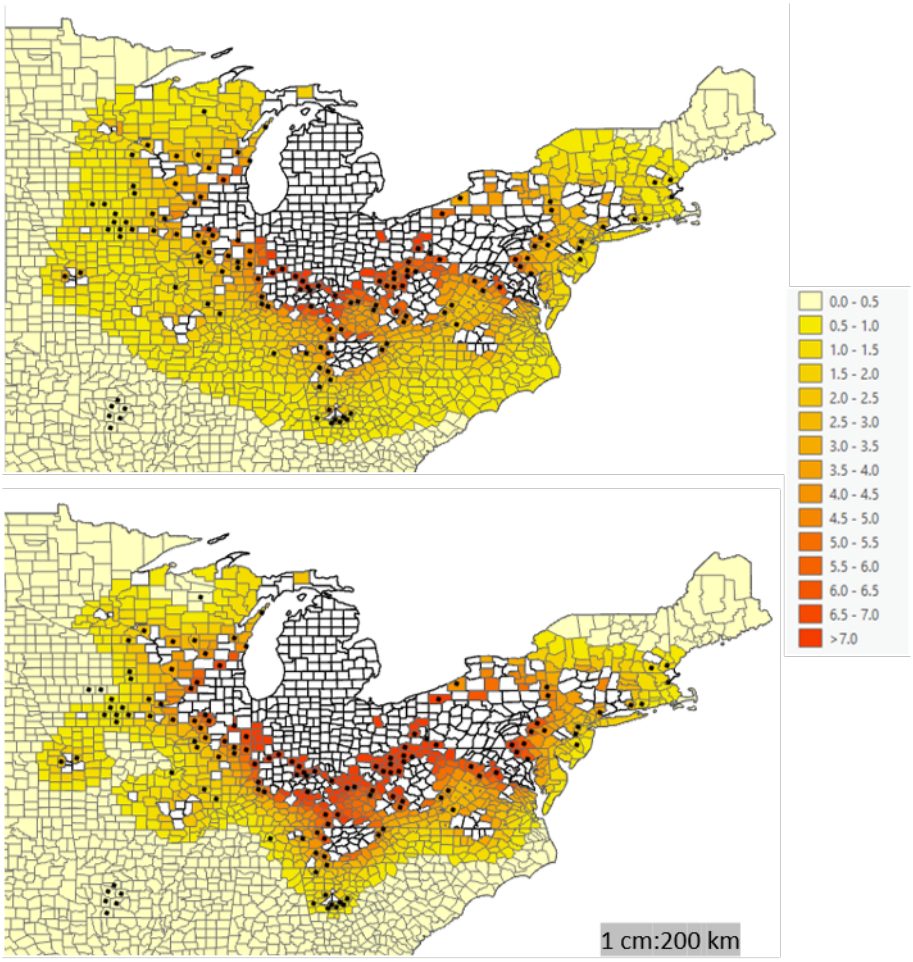


Figure 6.7: Power-law kernel function (upper) based on model 1 ($\exp^{A}_{\alpha_1} K_{pwr}$) and negative exponential kernel function (below) from model 5 ($\exp^{A}_{\alpha_1} K_{\exp}$) where both functions had been applied to estimate the spread of EAB in 2014. The λ value is 1.0315661 for power-law kernel and it represents the mean value of the estimated parameters by using model 1 ($\exp^{A}_{\alpha_1} K_{pwr}$) while for negative exponential formula, the λ value is 2.121683 which is the mean value of the estimated parameters by using model 5 ($\exp^{A}_{\alpha_1} K_{\exp}$). The white counties represents the colonised counties before 2014, while the colour counties represents the sum of infection pressure from neighbour colonised counties. The light cream colour with black dots represents FN, while without dots indicates TN

Therefore, it can be understood that the EAB tends to colonise the nearby counties unless the ash species and the suitable MAGDD are unavailable. It can also be observed in the Figure 6.1 and Figure 6.8 that the most colonised couties are found in clusters. However, there are some unexpected colonisation events. Further, Figure 6.8 displays boxplots of the dispersal distance (d_{ij}) within $d_{max} < 12$ degree between colonised counties on a particular year (T) and colonised counties that were colonised in all previous years ($2002, T-1$). Hence, the interquartile range (IQR) reveals the degree

of clustering, where a small IQR (small box of the boxplot) reflects majority of the colonisation in cluster, and dispersal distances are mostly within 8 degree over 18 years. The increasing distance over time is due to the fact that EAB is spreading over a wider area of the USA as time progresses. Hence, the box in the Figure 6.8 confirms the colonisation behaviour and level which aligns with the negative exponential kernel.

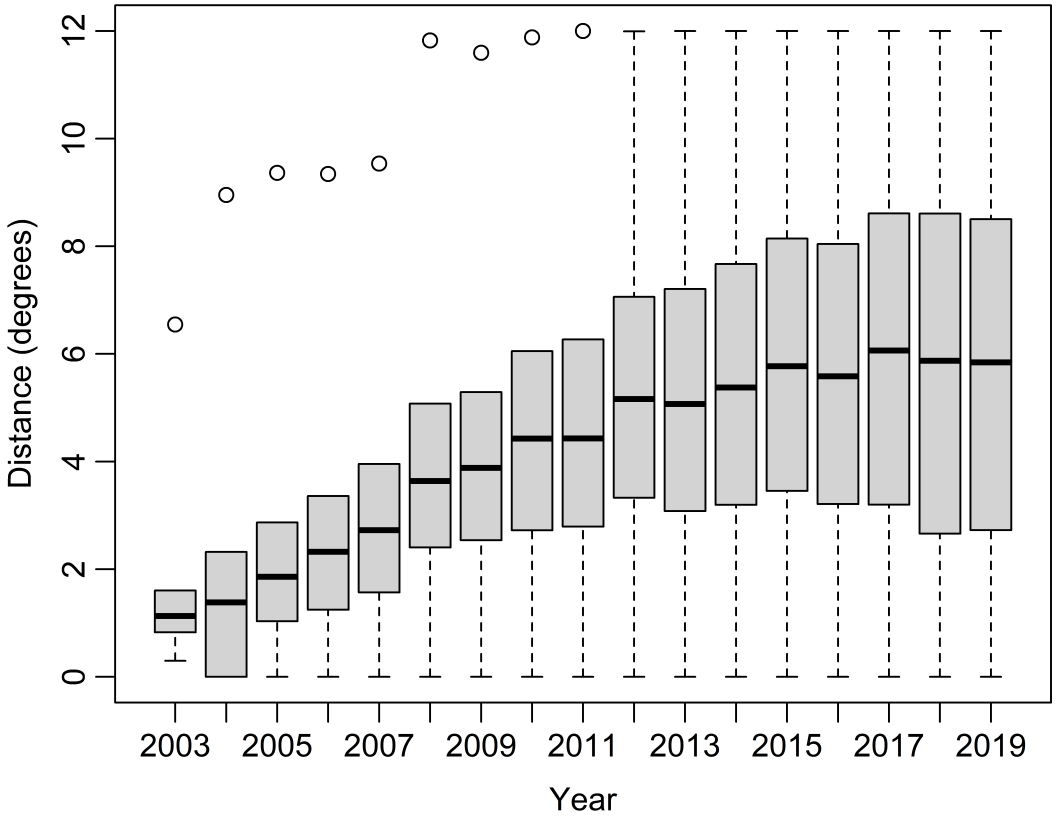


Figure 6.8: Boxplots of the dispersal distance between colonised counties on a particular year and colonised counties that were colonised in all previous years.

Comparison 2 and 3: Optimal scaling method and MAGDD function

A range of the MAGDD functions and scaling methods were applied in the all models in order to (i) minimise the false numbers and improve model performance, and (ii) understand the relationship between EAB distribution and the MAGDD. These functions were the exponential, normal distributions and gamma distributions. In addition, these models had been applied at different scaled for MAGDD. The four models which are Model 22 ($G^{C}_{\alpha_1,0.5}K_{\exp}$), Model 15 ($G^{E}_{\alpha_1,0.5}K_{\exp}$), Model 17 ($G^{D}_{\alpha_1,0.5}K_{\exp}$) and Model 21 ($G^{F}_{\alpha_1,0.5}K_{\exp}$) were identical in every aspect except for the scaling method used. Thus, to understand the impact of scaling on model effectiveness, it lead us to compare the

top-performing once, Model 15 ($G^{E}_{\alpha_1,0.5}K_{\exp}$), with the lowest-performing model which is Model 22 ($G^{C}_{\alpha_1,0.5}K_{\exp}$). These comparison highlight the crucial role of scaling in model performance, because there were a noticable variation in the amount of false numbers. The scaling method c^Eenhances to record lower false cases and improves the model ability through increasing the accuracy of the classifcation and events prediction. Hence, it is the optimal scaling method based on this research. In short, the scaling method can be a significant factor, influencing EAB colonisation and the c^E is the optimal one, which might be because of only estimating one parameter of the MAGDD function.

In addition, Model 24 ($G^{E}_{\alpha_1,1}K_{\exp}$) and Model 15 ($G^{E}_{\alpha_1,0.5}K_{\exp}$) were identical in every aspect except on their fixed shape parameters, where α_2 was 1 and 0.5, respectively. These models are also the second and the third best-performing models, respectivly. Hence, they can help to identify the optimal MAGDD function, and the relationship between heat availability and EAB colonisation behaviour. Although there were minor difference between these models, McNemar's test was used to compare performance of two models. The test p-value ($<$ 2.2e-16) reveals significantly different between the performance of the two models in terms of both sensitivity and specificity. Hence, it can conclude that the Model 24 ($G^{E}_{\alpha_1,1}K_{\exp}$) fits the data better, as it has a lower number of false events. In addition, a shape parameter of 1 in a gamma distribution indicate an exponential decay behaviour, where this model is the second best-performing model. This aligns with the findings of the best-performing model 5 ($\exp^{A}_{\alpha_1} K_{\exp}$) which shows exponential decay function. Therefore, based on the consistency of this decay pattern in the top-performing models, we conclude that the exponential function is the most suitable model.

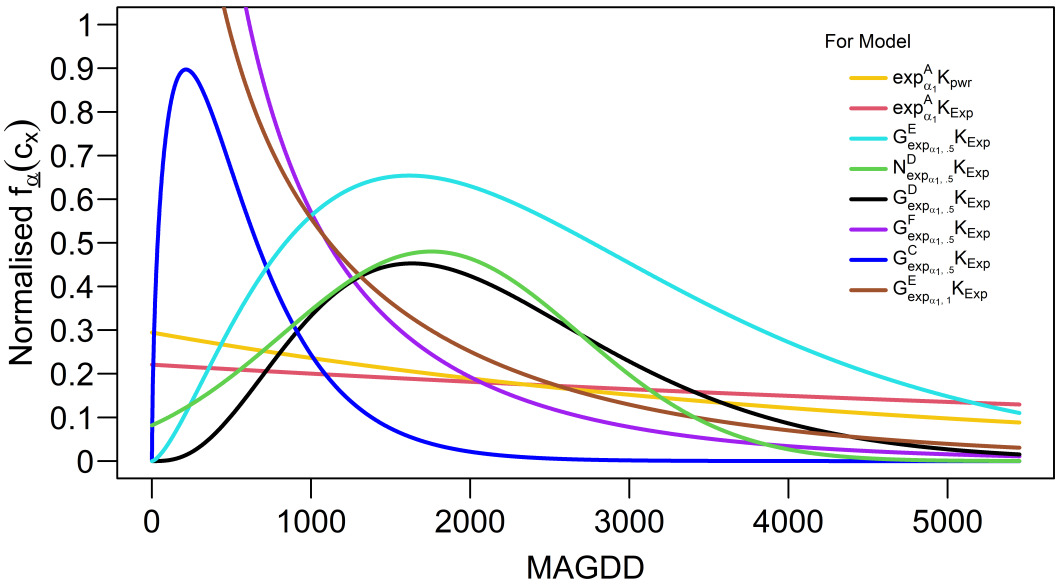


Figure 6.9: Normalised curves represent the MAGDD values across various models studied. Normalisation is conducted swhere the area under the curve equals one.

6.3.5 Additional Insights: Analysis of the suboptimal models

In light of our findings, it has been confirmed that the models 16 ($N^{D}_{\alpha_1,0.86}K_{\exp}$), 17 ($G^{D}_{\alpha_1,0.5}K_{\exp}$) and 22 ($G^{C}_{\alpha_1,0.5}K_{\exp}$) demonstrate higher overall false rates than others. The Table 6.9 confirmed this by revealing a higher overall false rates for these models. This is because the kernel curves of these models are relatively flat. Hence, they reflect the minimal impact on EAB movement (see Figure 6.6). Accordingly, the suitability s_i dominate these models behavior. It can be ssen that Figure B.15 in Appendix B.3 shows the suitability maps for 16 ($N^{D}_{\alpha_1,0.86}K_{\exp}$) and 17 ($G^{D}_{\alpha_1,0.5}K_{\exp}$) models. Both maps closely resemble the original data of ash intensity and MAGDD (see Figure 6.2b and Figure 6.3). This indicates that the influence of the kernel term is minimal. In addition, the suitability map of models 22 ($G^{C}_{\alpha_1,0.5}K_{\exp}$) reflects that the model fails to integrate all components (i.e. suitability and EAB dispersal). The map only shows the influence of the lowest MAGDD values.

To provide further clarification, the prediction maps for these models to predict the EAB colonisation counties in 2014 are given in Figure 6.10 and Figure 6.11. It is clear that all these three figures are indication that the suitability has more impact on the risk map than the kernel. For example, the darker area in the east in Figure 6.10 are the area where ash intensity and MAGDD are high. In addition, the impact of the kernel in

the clustered counties is not easily observable, with only a few noticeable effects. The prediction map in Figure 6.11 demonstrates with clarity the impact of the MAGDD lowest values in the top south region, while the outcomes of the suitability factors in the east.

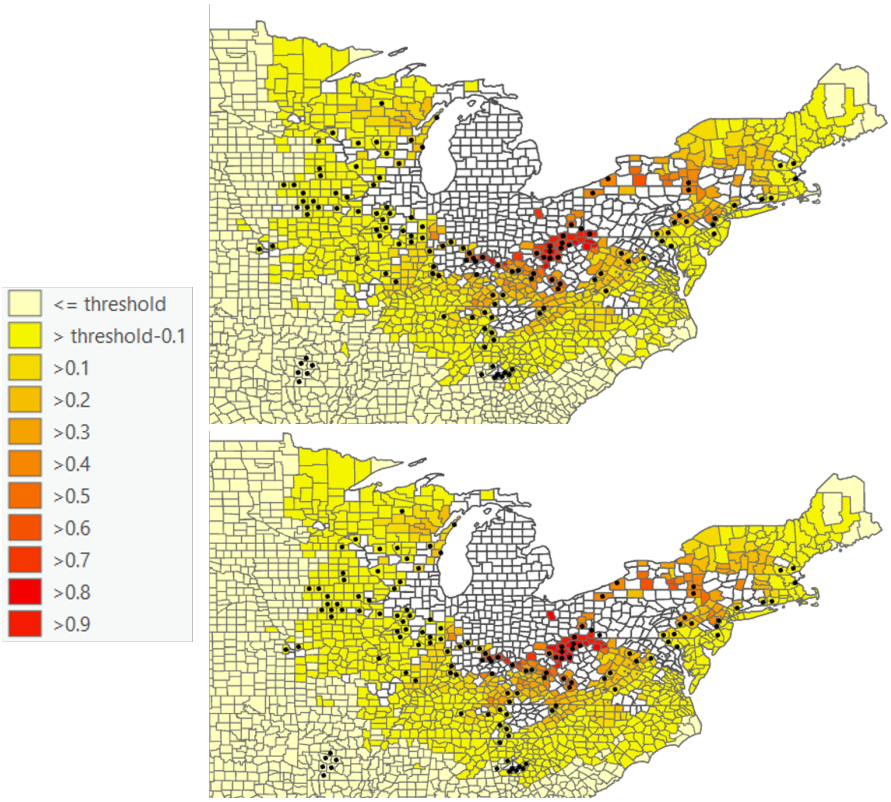


Figure 6.10: Model 16 ($N^{D}_{\alpha_1,0.86}K_{\exp}$) (upper) and Model 17 ($G^{D}_{\alpha_1,0.5}K_{\exp}$) (below) where both had been applied to predict the spread of EAB in 2014 based on the median estimation parameters for the model in observed data from 2002 to 2013. The white counties represents the colonised counties before 2014, while the colour counties represents the sum of infection pressure from negibour colonised counties. The black dots are EAB observations in 2014. The light cream colour with black dots represents FN, while without dots indicates TN.

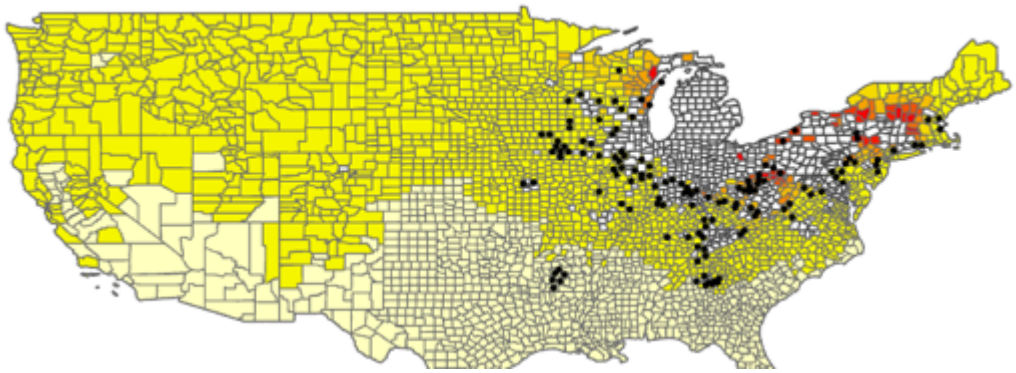


Figure 6.11: Model 22 ($G^{D}_{\alpha_1,0.5}K_{\exp}$) predict the spread of EAB in 2014 based on the median estimation parameters for the model in observed data from 2002 to 2013. The white counties represents the colonised counties before 2014, while the colour counties represents the sum of infection pressure from negibour colonised counties. The black dots are EAB observations in 2014. The light cream colour with black dots represents FN, while without dots indicates TN.

6.3.6 Best model is model 5 ($\exp^{A}_{\alpha_1} K_{\exp}$)

Model 5 ($\exp^{A}_{\alpha_1} K_{\exp}$) stands out as the best model performance. This model underscores the critical rule of the positive linearity of ash intensity and the negative exponential of MAGDD values to identify EAB colonisation counties. The model also reflects the ability of the negative exponential kernel function to capture EAB movement. The following analysis, supported by mapped visualisations, highlights the model's effectiveness.

The suitability map provides clear visual evidence that the model well captures the variability of the ash intensity and MAGDD factors over counties (see Figure 6.12). The map effectively identifies the counties with ash intensity, represents by the top two darker colours. Regarding the MAGDD factor, there is a clear negative relationship as the MAGDD values increase, the rate decreases. Accordingly, the model assigns high rates to MAGDD values ranging between zero to 2796. Moreover, the prediction maps for specific years namely 2004, 2007 and 2012, while it also can be provided to all other study period years, confirm the model's effectiveness in integrating suitability and dispersal movement factors (see Figure 6.13). Thus, the model can predict the TP counties with a higher rate for counties adjacent to clustered colonised counties. The model also enhances the effectivence of the kernel to illustrate EAB distribution, where EAB prefers to flow across nearvy counties.

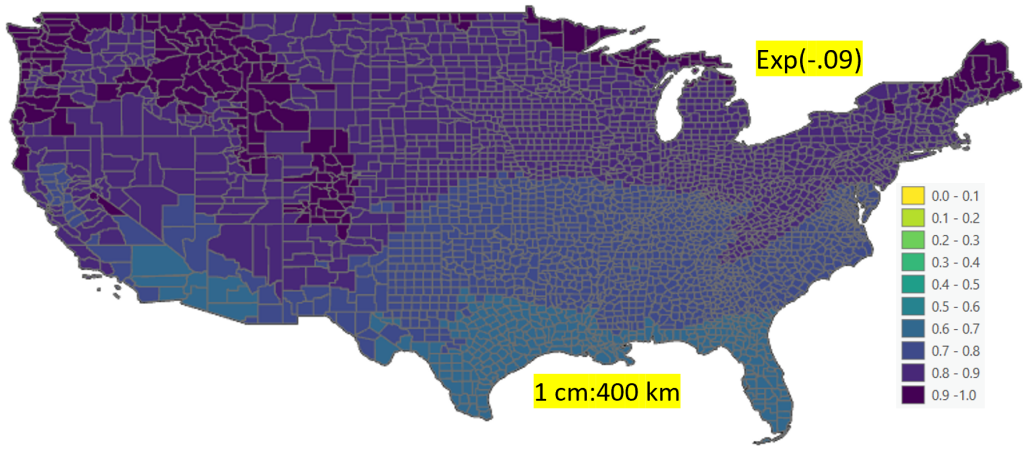


Figure 6.12: The overall suitability for each county in the USA for the invasive EAB colonisation under the model 5 ($\exp^{A}_{\alpha_1} K_{\exp}$). The value calculated for each county is the median of the posterior distribution of the suitability function 6.4.

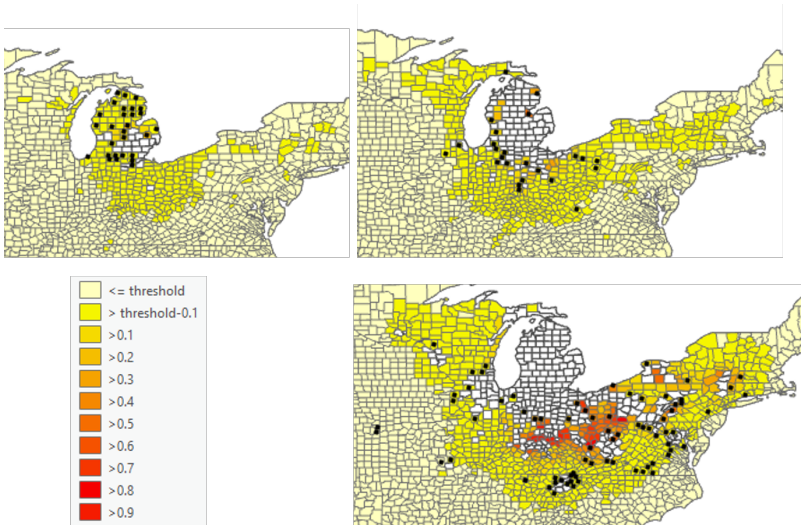


Figure 6.13: Model 5 to predict the spread of EAB in 2004 (top left), 2007 (top right) and 2012 (below) based on the median estimation parameters for the model in observed data from all previous years. The white counties represents the colonised counties before year T, while the different colour counties represents the sum of infection pressure from negibour colonised counties. The black dots are EAB observations in the prediction year (T+1). The light cream colour with black dots represents FN, while without dots indicates TN. The other colours reflect the positive counties, where with black dots mean TP and without indicates FP.

6.4 Summary and discussion

The Emerald Ash Borer (EAB) is a non-native insect, which has posed a significant threat to native ash species (*Fraxinus*) in the USA. However, it is a challenge to erad-

icate EAB, because of its biological characteristics. In essence, the EAB larvae develop inside the host ash wood tree for 1-2 years [97]. Efforts to combat EAB infestations have included accurate understanding of its dispersal and colonisation through statistical analysis. Hence, this chapter primarily aimed to statistically estimate the spatial and temporal spread of the EAB across USA counties. Thus, the model focused on the county unit on an annual temporal scale between 2002 when EAB was first identified in USA to the end of 2019.

This study, focused on generalisations of the generic colonisation–dispersal model proposed by Catterall et al. (2012) [103]. This is because that the approach allows for the inclusion of the invasive species distribution at multiple time points. In addition, it takes accounts of the temporal uncertainty in colonisation times. Further, it allows to include climate and non climate covariates in order to capture variation in the suitability of counties for invasive colonisation. Moreover, this approach applied MCMC to deal with model complexity due to multiple parameters and enables quantification of uncertainty in model parameters, inferred unknown colonisation times and prediction of future colonisations.

Different models were applied to adapt the model of Catterall et al. (2012), each of them included the same factors which are dispesal distance between counties and two covariates that vary spatially across counties, the mean of the annual growing degree day (MAGDD) and normalised ash intensity. The models varied due to several factors: (i) the exponential function, gamma distribution, and normal distribution were utilised to estimate the MAGDD variables; (ii) different scales were applied to the MAGDD variable; and (iii) two distinct kernels were used, namely the power-law kernel and the negative exponential kernel. The optimal performance model with the lowest number of incorrect classification counties was model 5 ($\exp^{A}_{\alpha_1} K_{\exp}$), while the worest performance model was model 22 ($G^{C}_{\alpha_{1,.5}} K_{\exp}$).

In addition, it is worth to mention that all models contributed viewpoints about EAB reaction with the given different factors and functions. For example, (i) the comparison between models ($\exp^{A}_{\alpha_1} K_{pow}$) and ($\exp^{A}_{\alpha_1} K_{\exp}$) provided that EAB dispersal behavior tends to be mostly very localised. Thus, (ii) the negative exponential kernel is the best fit. Further, the decay exponential function is the most effective option to reflect the relationship betwen EAB colonisation and MAGDD factor, as the best performance model ($\exp^{A}_{\alpha_1} K_{\exp}$) and the second-best model 24 ($G^{E}_{\alpha_{1,1}} K_{\exp}$) underscored

that. Furthermore, all models confirmed that (iii) there is a positive relationship between suitable county to EAB colonisation and normalised ash intensity.

6.4.1 Suitability and EAB

Orlova-Bienkowskaja et al. [97] demonstrated that the EAB adult is an unable to colonise an area while its MAGDD smaller than 700 degree. This figure is based on the observed EAB in America, Europe and Asia, where Asia is the native of the EAB. On the other hand, the suitability map (see Figure 6.12) of this study found a negative relationship between EAB presence and MAGDD, with the highest suitability for counties with the MAGDD values less than 700 degree-days, while the number of these counties are very low comparing with others. However, because these counties have lack normalised ash intensity (almost zero) and are far from the nearest colonised counties, model 5 ($\exp^{A}_{\alpha_1} K_{\exp}$) did not predicte EAB in these counties. Thus, counties with less than 700 degree-days, they did not impact the model's performance. Hence, to improve our understanding, more data on EAB presence in regions with MAGDD values below 700 degree-days is needed.

Furthermore, the body size of the EAB females in the colder areas (lower MAGDD) are being smaller [223]. In addition, these females might have difficulty for surviving and cannot lay enough eggs [223, 97]. On the other hand, the study conducted by Haack et al. (2022) [224] emphasised that adults and larvae of the EAB complete mortality when they expose to 56°C or above for at least 30 continuous minutes duration. Thus, this finding suggests that the increasing of the heat availability (high MAGDD) in an area may impact the reproduction and persistence of the EAB adult. Hence, it is unfavorable for the EAB colonisation.

6.4.2 EAB dispersal mechanisms

Moreover, the study outcomes revealed that including the geographical distance between the center of the current colonised counties with the nearby non colonised counties is an essential factor to understand EAB. This can be supported with models ($N^{D}_{\alpha_1,0.86}K_{\exp}$), ($G^{D}_{\alpha_1,0.5}K_{\exp}$) and ($G^{C}_{\alpha_1,0.5}K_{\exp}$) which have flat kernels and the highest false rates comparing with other models. In addition, useing a negative exponential kernel than a power law kernel was more suitable model to understand the EAB colonisation dynamic. Hence, it increased the model effectivness and performence. Further, compared

to the power law kernel, the negative exponential kernel reveals a sharp drop as the distance rises. This is aligned with the EAB colonisation behaviour, where the colonised counties tended to form clusters, with some odd (see Figure 6.1 and Figure 6.8). Accordingly, the natural migration of the EAB adults tend to be local distance dispersal as long as the ash species and the suitable MAGDD are available.

Furthermore, Muirhead et al. (2005) [43] demonstrated the effectiveness of the negative exponential kernel to predect the invaded locations with EAB in 2004 within the Great Lakes state. They found that, even at low probabilities of dispersal, the model predict correctly 77 percent the of invaded locations. In addition, Muirhead et al. (2005) confirmed that the majority of the EAB infestations between 2002 to 2005 occured because of the natural EAB adult flights. Hence, the negative exponential kernel can accurately simulate the EAB colonisation distribution, particularly in the natural dispersal of the EAB adults. On the other hand, human-mediated activities such as the movement of infested seedlings, can lead to the long-distance dispersal [99], causing unexpected EAB infestation. For example, in 2003, an outbreak of the EAB reached Maryland state, USA from Michigan state, USA [99]. Moreover, a 600 miles from the nearest infested area, EAB outbreak was discovered in Colorado state, USA in 2013 [99]. Thus, the localised spread of EAB should have significant implications for management strategies. The priority can present for colonised counties and where it is projected to spread next. Further, it helps researcher to conduct and focus a research in specific areas, where they can gather valuable data and enhance the understanding of the EAB.

Chapter 7

Conclusion

7.1 Overview

The escalating global population and the increasing interconnectivity of worldwide systems pose substantial challenges for the sustainability of food production and the preservation of plant forests. One of the primary concerns is the rise of non-native insect pests that threaten crop yields and plant health. Hence, they are leading to reduce crop yields, which cause food insecurity. Moreover, they create a severe threat to the productivity and diversity of native ecosystems, and significant financial losses for industries, property owners, farmers and governments.

Therefore, accurate identification of pests in the proper physical location and time is crucial for effective management techniques and understanding of their population dynamics. Conventional pest identification approaches mainly depend on specialist knowledge, which are frequently limited and costly. Hence, experts such as entomologists are unable to effectively and quickly monitor wider spatial space. On the other hand, creative methods that integrate artificial intelligence (AI) with citizen science can provide more comprehensive monitoring of geographical area and earlier detection of pest invasions. However, these modern methods may have lower accuracy compared to traditional methods.

Lower accuracy of the citizen science AI method or limited experts advice can cause direct adverse impacts for farmers, foresters and researchers. Farmers and foresters can overuse or misuse pesticides, which can harm non-target species. Researchers often need to invest additional resources to establish a gold standard for data validation, ensuring the accurate identification of AI tools, which can lead to increased project

costs. Limited availability of expert data can negatively effect the research findings, increasing the likelihood of biased results.

This thesis demonstrates its strength by developing novel statistical approaches for analysing and modelling tow non-native pests using two different types of datasets: (i) a citizen science AI dataset with absent gold standard, and (ii) a limited expert dataset. The two non-native pests are (i) Fall Armyworm (FAW, *Spodoptera frugiperda*) in maize crops (*Zea mays*) in India during the period 2018-19, and (ii) the Emerald Ash Borer (EAB,*Agrilus planipennis*) epidemic in ash trees (*Fraxinus*) in the USA from 2002 to 2020. The FAW data are based on a citizen science approach utilising Plantix, an innovative method that integrates artificial intelligence (AI) with mobile technologies. Constructed by PEAT GmbH, it can provide comprehensive monitoring of geographical areas and early detection of pest invasions. For the EAB, it is an expert dataset, which was collected by USDA APHISUS. It includes an initial true positive case from each observed infested county in the USA. Statistical analysis in both dataset has different objectives. For the citizen science AI dataset, the estimation of the gold standard is at first, and it is before investigating pests dynamics. For the expert dataset, we typically focus on understanding the population dynamics of pests.

The decision to focus on FAW and maize was not random; rather, there were reasons. In India, maize is the third staple crop after rice and wheat, and FAW was first detected in 2018 on the maize crop. Hence, FAW has caused widespread concern as it leads to sever damage. This is because it has a high reproductive rate and strong dispersal capabilities. The EAB, on the other hand, is one of the most destructive invasive forest pests found in the ash species of North America, causing widespread mortality of ash trees in urban landscapes and natural forests. It has killed hundreds of millions of ash trees in North America and is considered a high-impact pest. This has resulted in hundreds of millions of dollars being lost by municipalities, property owners, nursery operators, and the forest products industry. Controlling and eradicating EAB can be expensive, and have adverse effects on the environment. Thus, using statistical analysis can be an effective and budget-friendly manner to understand the distribution of EAB in order to control it.

7.2 Key contributions

The thesis offers vital contributions, some of which can be considered novel methodological contributions, which are valuable for scientific researchers. The contributions also benefit a wide range of stakeholders, and have a significant impact on the agricultural landscape. The findings of the statistical analysis of the citizen science AI dataset, specifically the Plantix app data in detecting Fall Armyworm (FAW) in India maize, contribute important knowledge to the PEAT company and any AI Developer. Moreover, these findings also serve farmers, entomologists and governments. In addition, the outcomes of the statistical analysis of the expert data to analyse the EAB in the USA extend the thesis contributions to the forestry sector.

Chapter 3 develops the first statistical framework for estimating True and False observations/models for citizen science AI dataset. A bi-normal mixture model is employed as the classification method. The latent variable of the mixture model presents and predicts True and False notifications, while we use our assumptions to classify the data into positive (unhealthy, pariculary in our data FAW) and negative (healthy, pariculary in our others than FAW) outcomes. The previous research used the latent variable to estimate positive and negative outcomes, thus, this approach is the first to use a classification model for estimating True and False observations. Following this, understanding the FAW intensity on maize over space (North and South India), seasons (Kharif and Rabi) and time (2018-2019) is estimated using *odds* metric, and it is the first contribution to studying the intensity of Fall Armyworm (FAW) over maize seasons, across India, and over time.

Chapter 3 provides a statistical method for distinguishing between True and False notifications, rather than relying on a cut-off that may not always represent a valid hypothesis. The findings from the *odds* calculation indicate that the invasion began earlier in 2018, before it was officially documented in mid-2018, which highlights the urgency to create a direct connections between warning app notifications and government as well as entomologists. This allows for early pest and disease management and prevent widespread outbreaks. Further, the presence of False notifications, especially more than 50 percent of the data are false, along with the non-interpretability of the *odds* curve of FAW intensity when including non-FAW data, underscores the need to determine accuracy with distinguishing between Healthy and non-FAW categories. This discussion leads us to **Chapter 4**.

Chapter 4 introduces a statistical approach to estimate the sensitivity and specificity of the citizen science AI dataset. This approach is the first to estimate pooled sensitivity and specificity within the same study, yet across distinct locations, using Bayesian meta-analysis. Each state assumes to be an independent study within the dataset of FAW in India. The approach also assumes that the logit sensitivity and specificity follow a multivariate normal distribution.

Chapter 4 findings suggest that FAW detection must be validated against Healthy reports rather than non-FAW reports, as there are valid biological reasons. In addition, the Plantix app has a sensitivity in detecting the FAW when Healthy as negatives, with approximately 50 percent of true positive cases being correctly identified. For the specificity, the Plantix app had the high specificity of 0.769 (0.730, 0.806), indicating its strong ability to identify Healthy maize crops, while around 20 percent of Healthy crops were misdiagnosied as infested with FAW (false positives).

These findings can inform the PEAT company to enhance the app accuracy and reliability in identifying specific pests or diseases in particular crops. This can be accomplished through improving the deep neural network (Plantix-DNN) algorithm. Following that, the PEAT can continuously collaborate with statisticians to ensure the app improvement. Further, these contributions not only serve the PEAT, yet also the users. The users use the app to identify crop issue, pay treatement and consider the app' advice for effective control. The incorrect diagnosis can lead to issues such as harming the tested crop, as the user may apply the incorrectly advised treatment. Further, since the app is imperfect, users should conduct independent verification to confirm the app diagnostic results and treatment advice. For example, users can conduct further research on the issue through online Resources. An alternative, users can just capture another photo for the affected crop, in order to ensure whether or not the app remains consistent. After esetimating sensitivity and specificity, it is time to move to one of the main objectives of the thesis: understanding FAW prevalence. This takes us to **Chapter 5**.

Chapter 5 estimates the prevalence of FAW across each state and within each season, applying various methodologies based on frequency analysis, Bayesian meta-analysis or the Bayesian meta-analysis with stochastic sensitivity and specificity, in order to select the most appropriate model. This study is the first study to estimate the spatio-temporal true prevalence of FAW in maize crop where the spatial scale was

state-level in India and the time term was the maize season. **Chapter 5** confirms that estimating FAW prevalence varies over time and space, which enhances the understanding of the preferred seasons and the risk states for FAW. The estimating sensitivity and specificity by season, rather than for the entire dataset, highlights that these metrics vary across different seasons. The lowest sensitivity value was recorded in Kharif 2018, at around 0.4, while the maximum sensitivity reached approximately 0.7 in Rabi 2019. Sensitivity tends to be higher during the Rabi season compared to the Kharif season, suggesting that FAW may be more prevalent and damaging during Rabi. These findings involve entomologists and governments in a deeper understanding of FAW behavior in maize fields and in implementing more effective pest management strategies during the preferred seasons and in high-risk states. Furthermore, the identified Rabi season as preferred for FAW encourages researchers to conduct more studies to understand the climatic factors influencing its increase. Additionally, during the Kharif season, FAW may prefer to infest other crops that are not planted during Rabi.

Further, the findings from the analysis of the citizen science AI dataset, specifically the Plantix app dataset, indicate that citizen science apps are very useful for reporting pest infestations. They allow threats to be detected rapidly, with high temporal resolution and wide spatial coverage of pest distribution. For statistical analysis, it is essential to include both positive and negative notifications, which are mostly missing in the expert data. These readings ensures self-sufficiency of the dataset, where these observations help estimate the intensity (**Chapter 3**), accuracy (**Chapter 4**), and prevalence (**Chapter 5**) of the subject of interest.

Chapter 6 develops the first a continental-level model for the spread of EAB in the USA over two decades. A colonisation-dispersal model is utilised to include climatic (annual average of growing degree day, MAGDD), non-climatic (ash intensity habitat) conditions, and dispersal mechanisms. The model is fitted to the best available data, quantifying the uncertainty in the model and its predictions and assessing its performance in tracking the spread of EAB over two decades, 2002 to 2019. EAB data consists of a single point from each county in the USA, therefore, including additional dataset help achieve the research objectivies related to understanding EAB colonisation behaviour.

Chapter 6 findings underscores that the significance of incorporating climatic and non-climatic factors as well as dispersal mechanisms. Chapter 6 concludes that there is

a positive relationship between suitable counties for EAB colonisation and normalised ash intensity. The decay exponential function effectively reflects the relationship betwen EAB colonisation and MAGDD factor. However, due to the limited data size, the research could not determine the exact range of suitable heat. The EAB dispersal behavior tends to be mostly very localised, making the negative exponential kernel the best fit. These findings assist various stakeholders in society in participating in the mitigation and control of EAB. Understanding that the EAB typically colonises locally, enhances the governments to develop management plans for neighboring uncolonised counties. Further, the analysis of expert data indicates that the overestimations findings may result from insufficient data in counties that are close to the colonized ones. Therefore, governments should find ways to involve citizen scientists in data collection. The following section will discuss some proposed future research directions based on the contributions of this thesis.

7.3 Research limitations

It is important to note the limitations of the study:

1. Time Recording Accuracy: Study findings may be biased or inaccurate if the collection data time is inexact. Experts and citizen scientists cannot guarantee to record observations at the right time or may miss some observations due to their availability in the infested area or fail to recognize the infestation.

2. Limited spatial coverage: The Plantix app data collection method may have focused on specific regions or areas based on the availability of the internet and smart phones, as well as farmers' skills and knowledge. Data from surveys also have limitations due to sampling locations or survey cost, as human experts cannot cover all infested areas. As a result, the study findings not being representative of the actual situation.

3. Generalisability: The findings and conclusions presented in this study could be unique to the study areas in India and the USA and would not be applicable to other geographical areas.

7.4 Future directions

For future research, this study suggests the methodological framework conducted for Fall Armyworm, can replicate on other pests affecting Indian maize or different crops in different space and time from the Plantix app dataset. One possible insect to analyse, it can be a violet stem borer, where it was the highest pest was found in the maize after FAW based on our Plantix app data used in this study. Additionally, repeating the thesis analysis for the latest version of deep neural networks (DNN-Plantix) could provide valuable insights about the improvement of the app accuracy. While smartphone-based citizen science tools for plant disease and insect pest detection using AI are increasingly popular, there are lack studies that conduct rigorous statistical analysis for assessing accuracy and reliability of these tools and further analysis. Hence, the statistical methodological frameworks in this research can be transferable to these tools as well or other that designed for use by citizen scientists.

Further, studing the prevalence of FAW in India during the epidemic phase can be as a baseline for comparison with more recent data. This can help to evaluate the changing patterns of the prevalence over time, and determining if the risk level has decreased. Moreover, it offers an approach to assess the effectiveness of the control strategies in managing FAW infestations. Furthermore, it can incorporate environmental factors into the Bayesian meta-analysis model in Method-4 in Chapter 5 to enhace the estimation of the FAW prevalence. Fan et al. [225] conclude that altitude (in meter), temperature (in degree Celsius), and humidity (in percentage) can significantly impact the survival, presence, reproduction and colonisation patterns of FAW. Therefore, these factors can be incorporated into the prevalence formula. Instead of using $\pi_i \sim beta(1,1)$, the following formula can be used, where the prior distributions for the parameters $\alpha_0, \alpha_1, \alpha_2, \alpha_3$ can assume to follow uniform.

$$logit\left(\pi_i\right) = \alpha_0 + \alpha_1 \times \text{temperature}_i + \alpha_2 \times \text{altitude}_i + \alpha_3 \times \text{humidity}_i.$$

The data for the these factors can be extracted with the same methodology as did for the EAB factors. Additional factors can also be included. However, it is crucial to confirm that each factor is significant before include in the model. One method to assess significance is by using the modified Gaussian model, as considered by Fan et al.

[225].

For the EAB, a limitation of this work is that the model used data from all previous years to predict the colonisation of the following year. Hence, this approach assumed that once a county was colonised, it remains permanently colonised (SI model). However, some infested ash died within 2 to 6 years after infestation [226, 227]. Moreover, a number of control actions, such as cutting infested hosts or using pesticides, were implemented. Hence, these actions can cause county to be categorised as in recovery rather than colonised. To address this limitation in future and enhance the model validity, it is possible to consider a shorter timeframe, for example every 6 years independently. This duration my be convenient, because the longest expected lifespan is 6 years.

Appendix A

Additional materials for Chapter 5

A.1 MCMC trace plots for Method-4

A.1.1 Trace plots for Method-4 during Kharif 2018

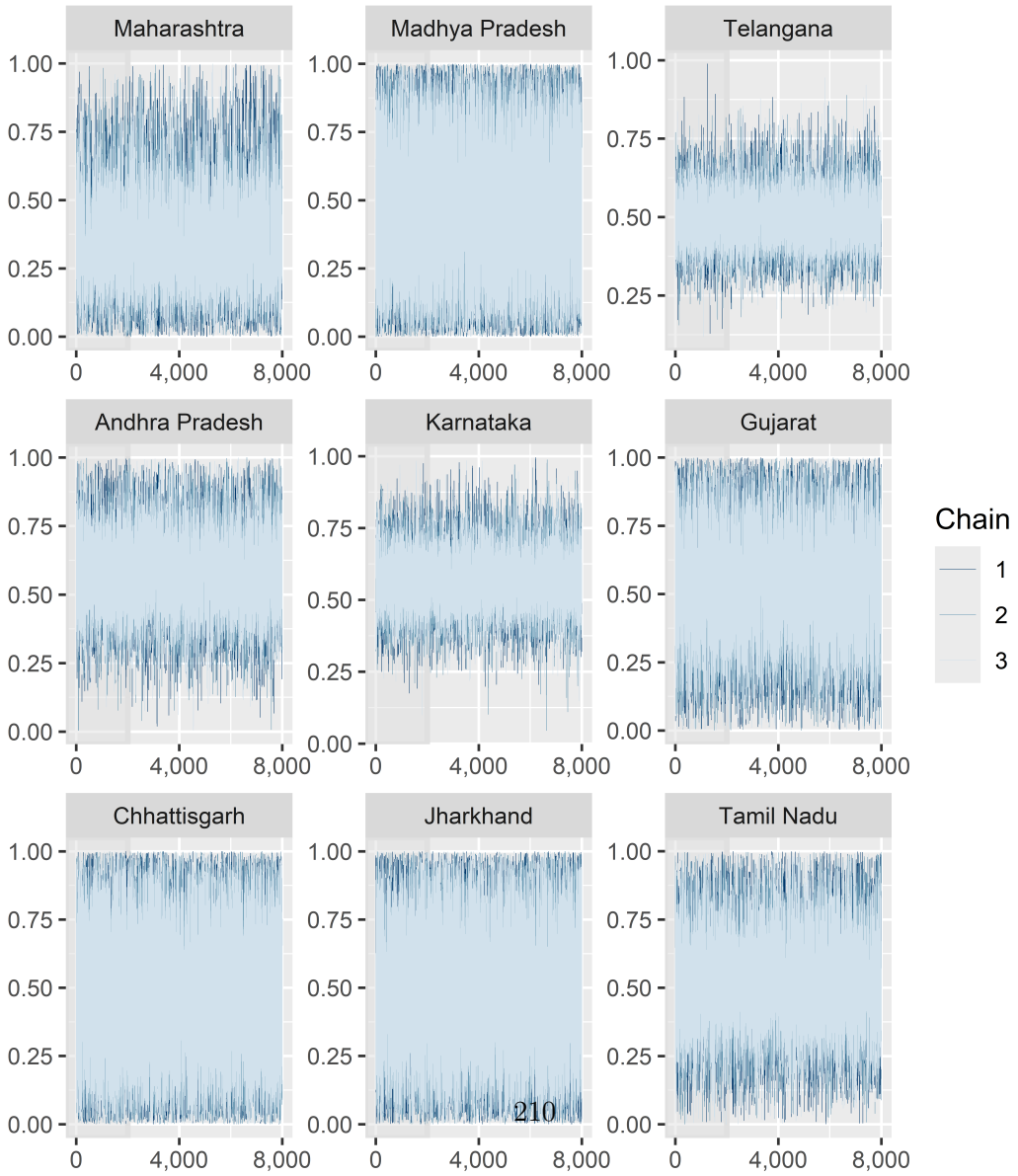


Figure A.1: Trace plots for the prevalence parameter during Kharif 2018, displaying three MCMC chains for each state using Method-4.

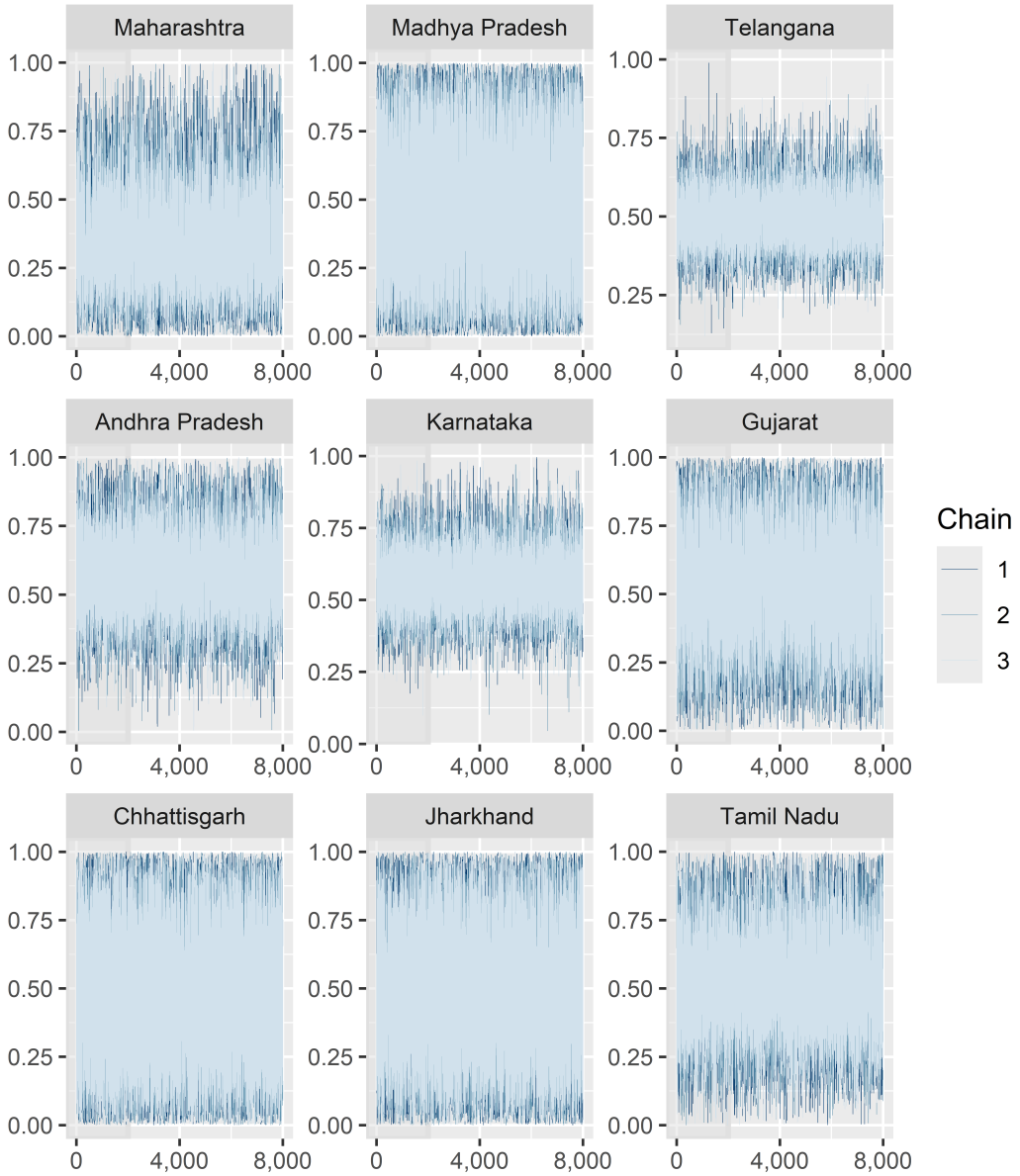


Figure A.2: Trace plots for the prevalence parameter during Kharif 2018, displaying three MCMC chains for each state using Method-4 after scaling the sample size by a factor of 10,000.

A.1.2 Trace plots for Method-4 during Rabi 2018

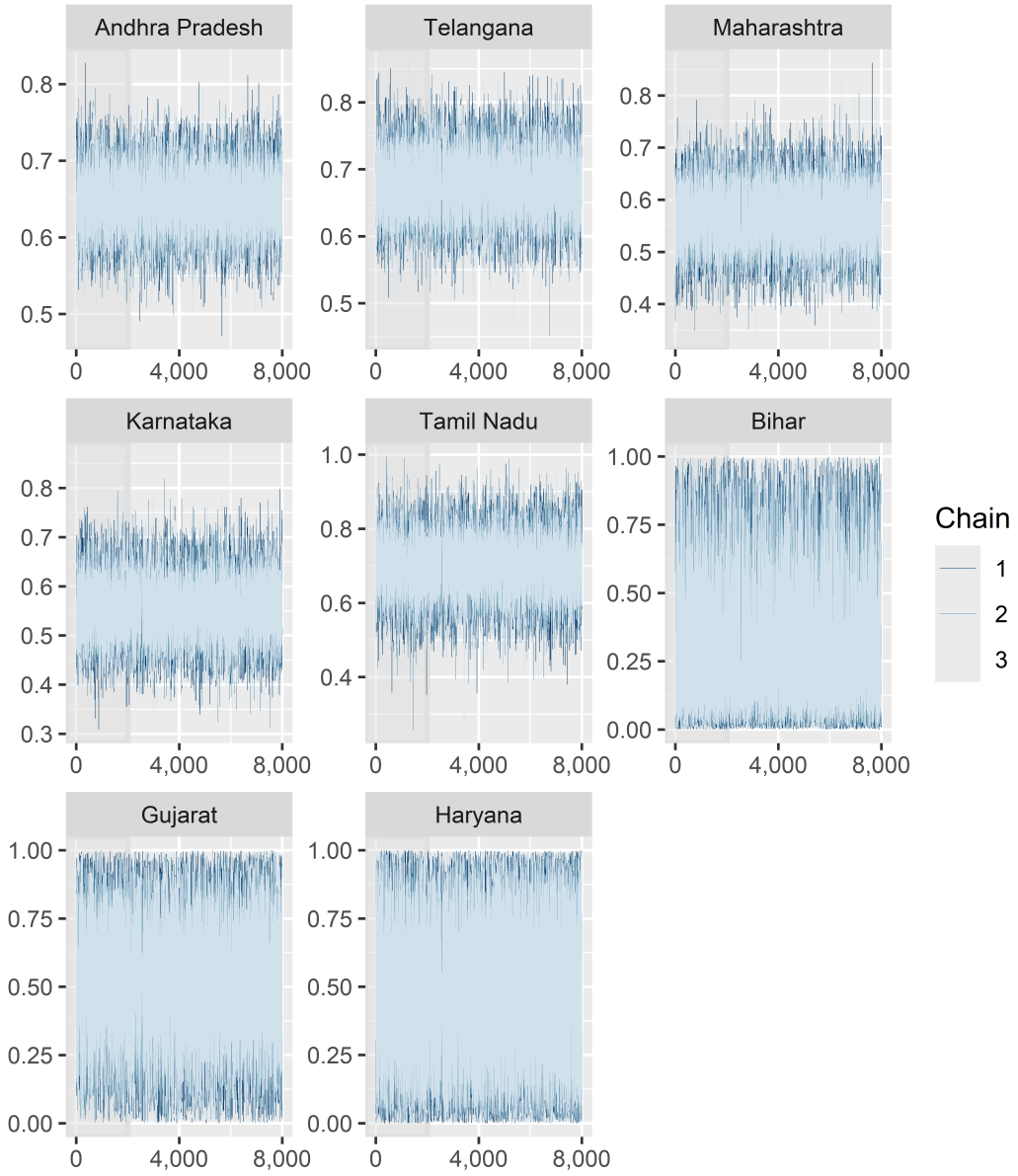


Figure A.3: Trace plots for the prevalence parameter during Rabi 2018, displaying three MCMC chains for each state using Method-4.

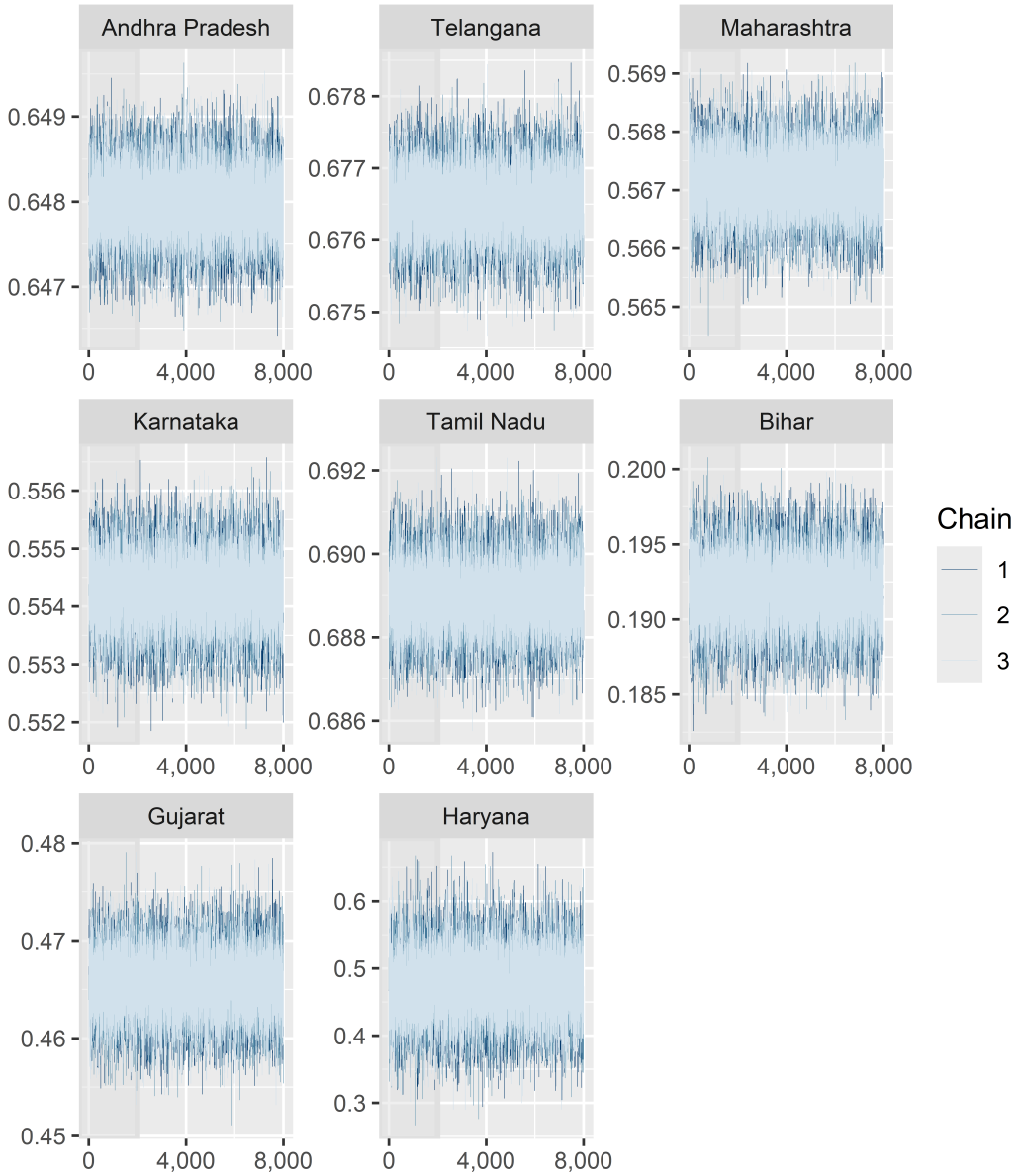


Figure A.4: Trace plots for the prevalence parameter during Rabi 2018, displaying three MCMC chains for each state using Method-4 after scaling the sample size by a factor of 10,000.

A.1.3 Trace plots for Method-4 during Kharif 2019

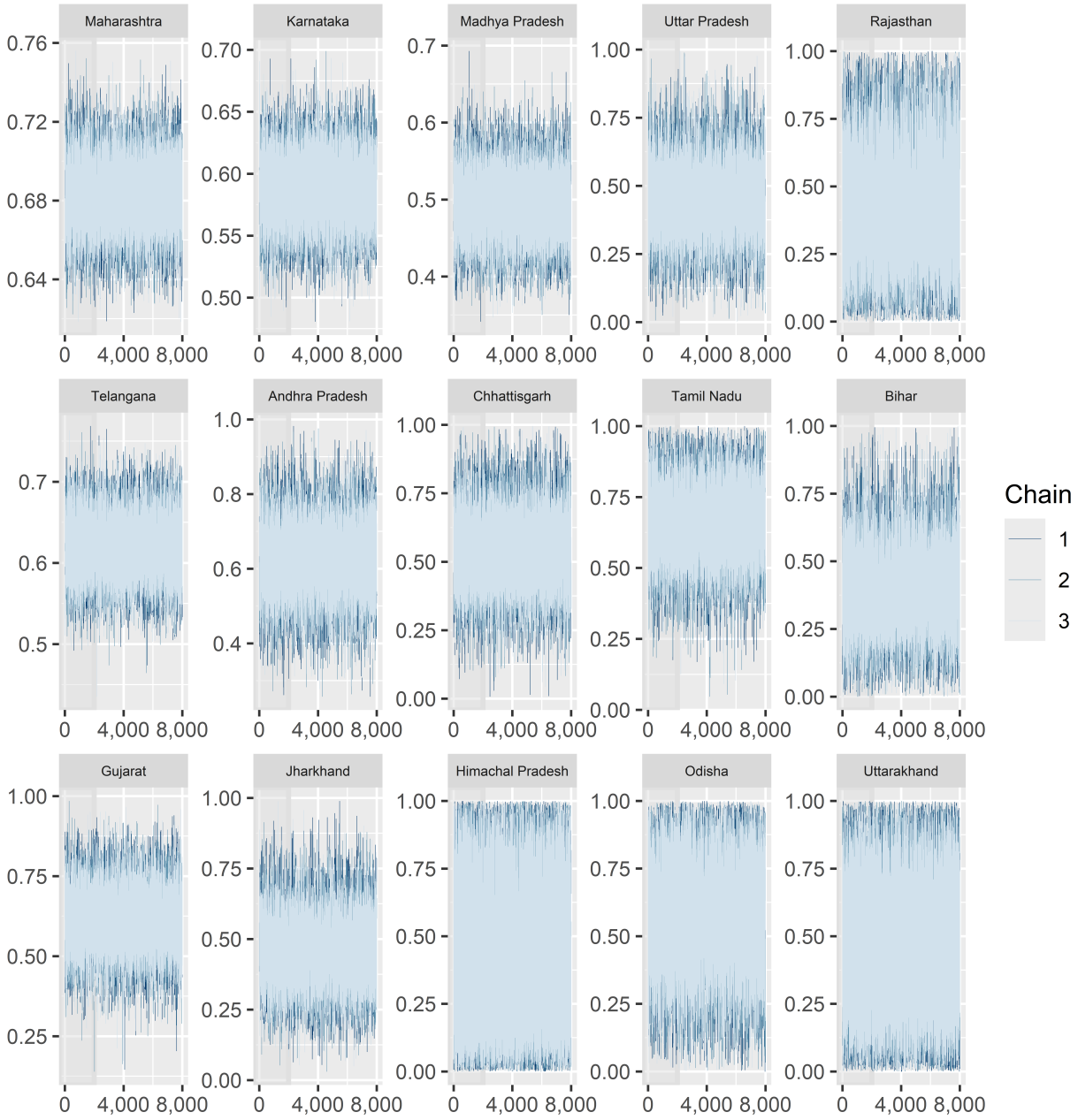


Figure A.5: Trace plots for the prevalence parameter during Kharif 2019, displaying three MCMC chains for each state using Method-4.

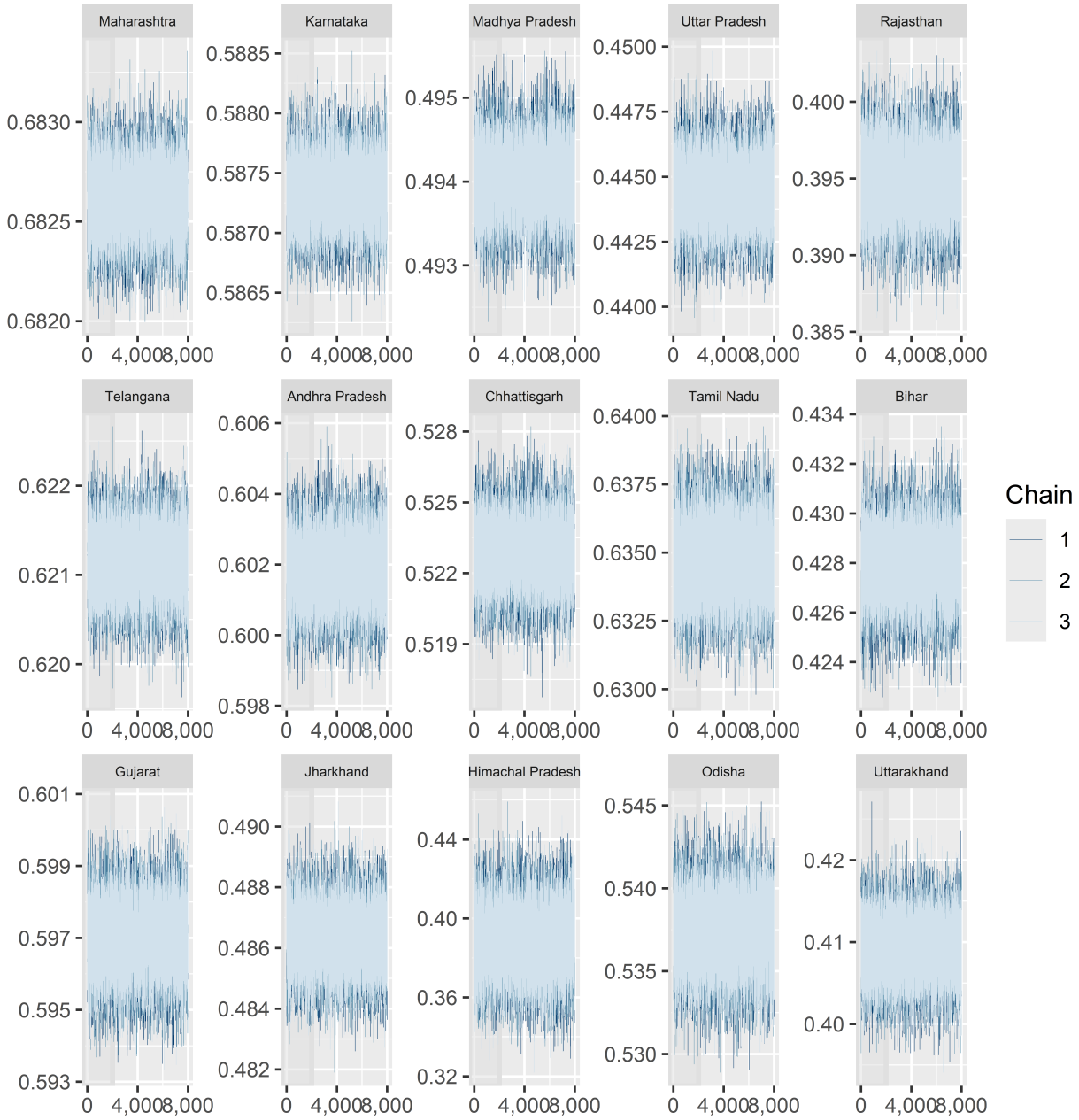


Figure A.6: Trace plots for the prevalence parameter during Kharif 2019, displaying three MCMC chains for each state using Method-4 after scaling the sample size by a factor of 10,000.

A.1.4 Trace plots for Method-4 during Rabi 2019

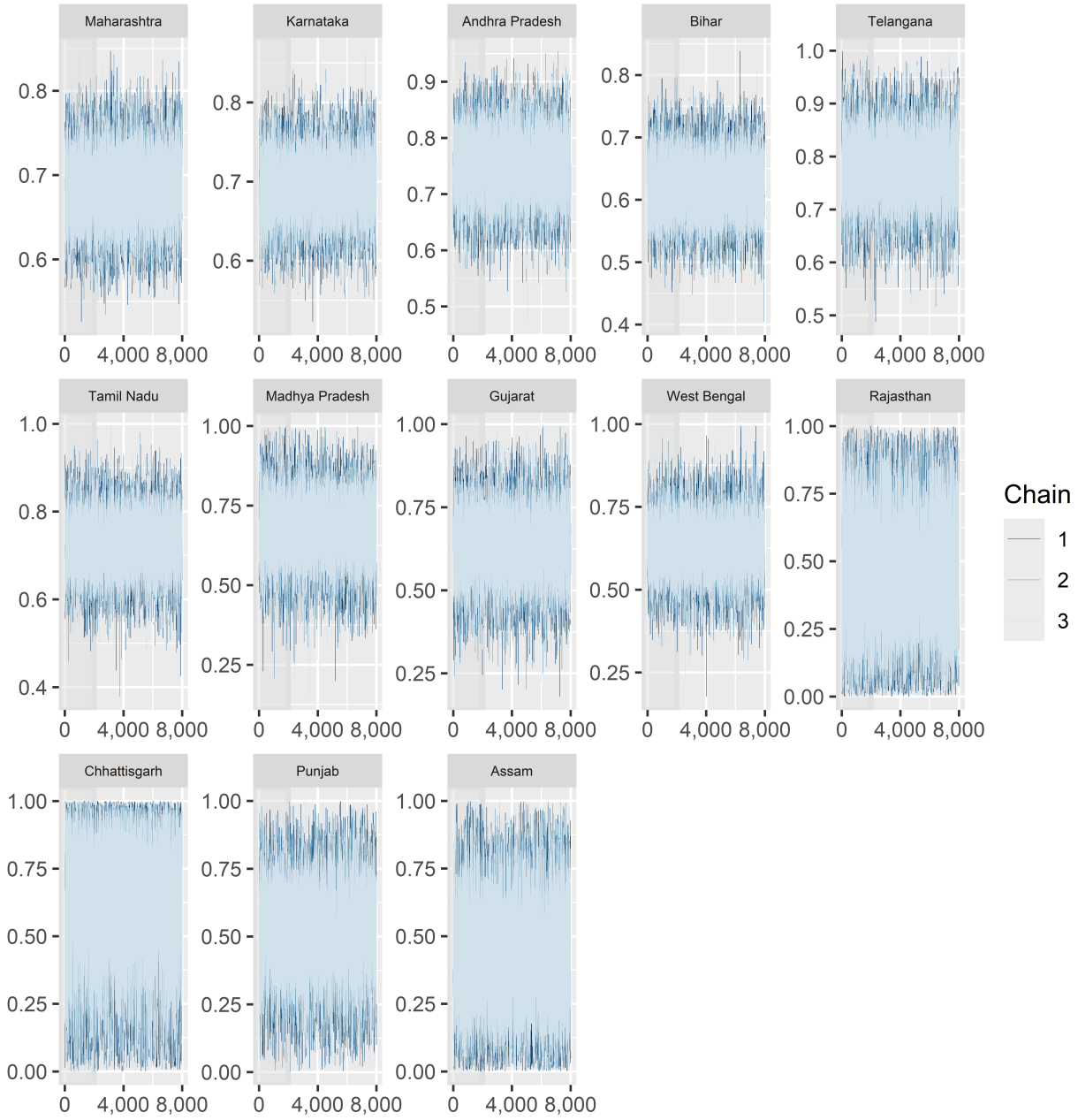


Figure A.7: Trace plots for the prevalence parameter during Kharif 2018, displaying three MCMC chains for each state using Method-4.

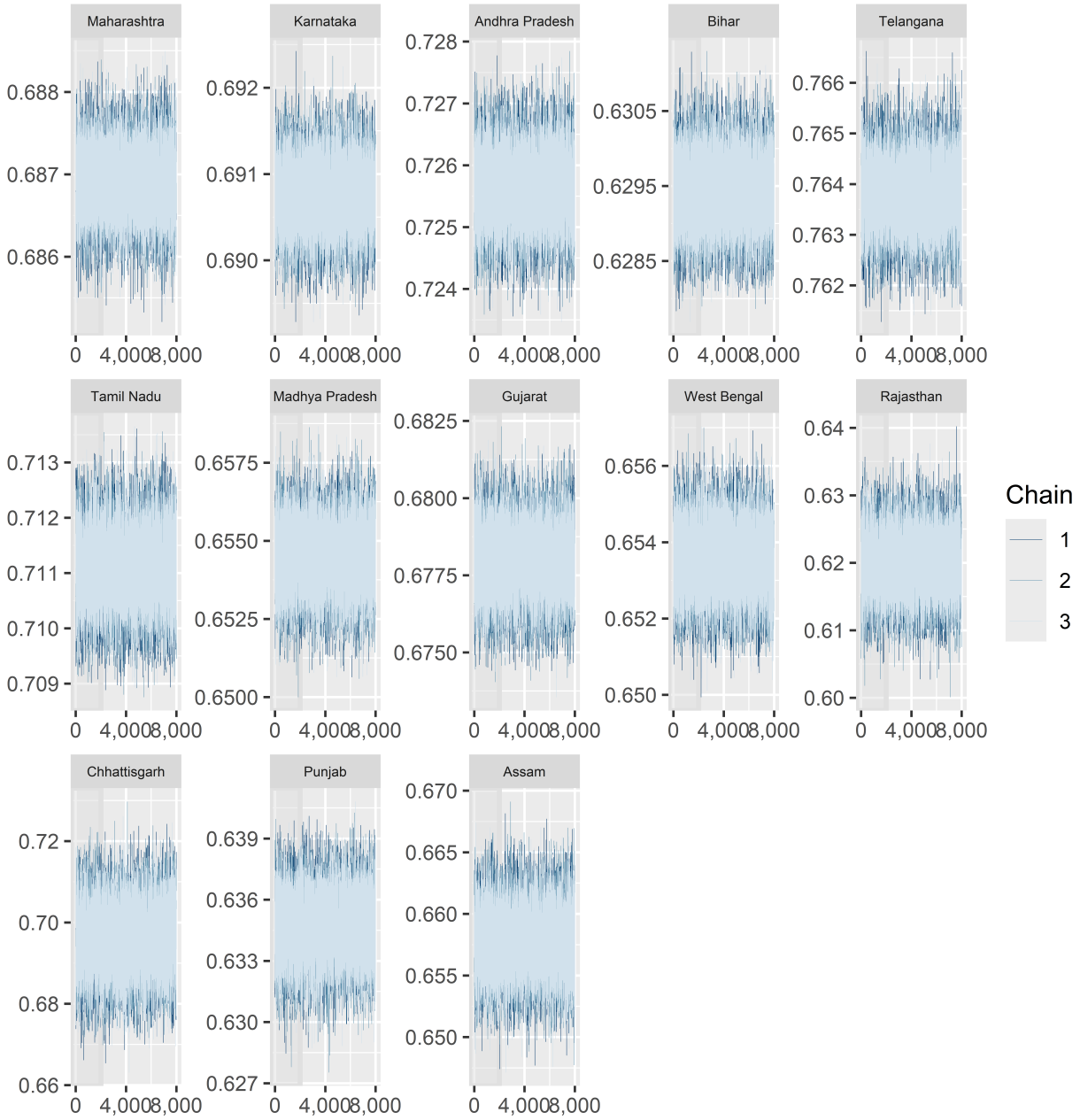


Figure A.8: Trace plots for the prevalence parameter during Rabi 2019, displaying three MCMC chains for each state using Method-4 after scaling the sample size by a factor of 10,000.

Appendix B

Additional materials for Chapter 6

B.1 MCMC convergence plots

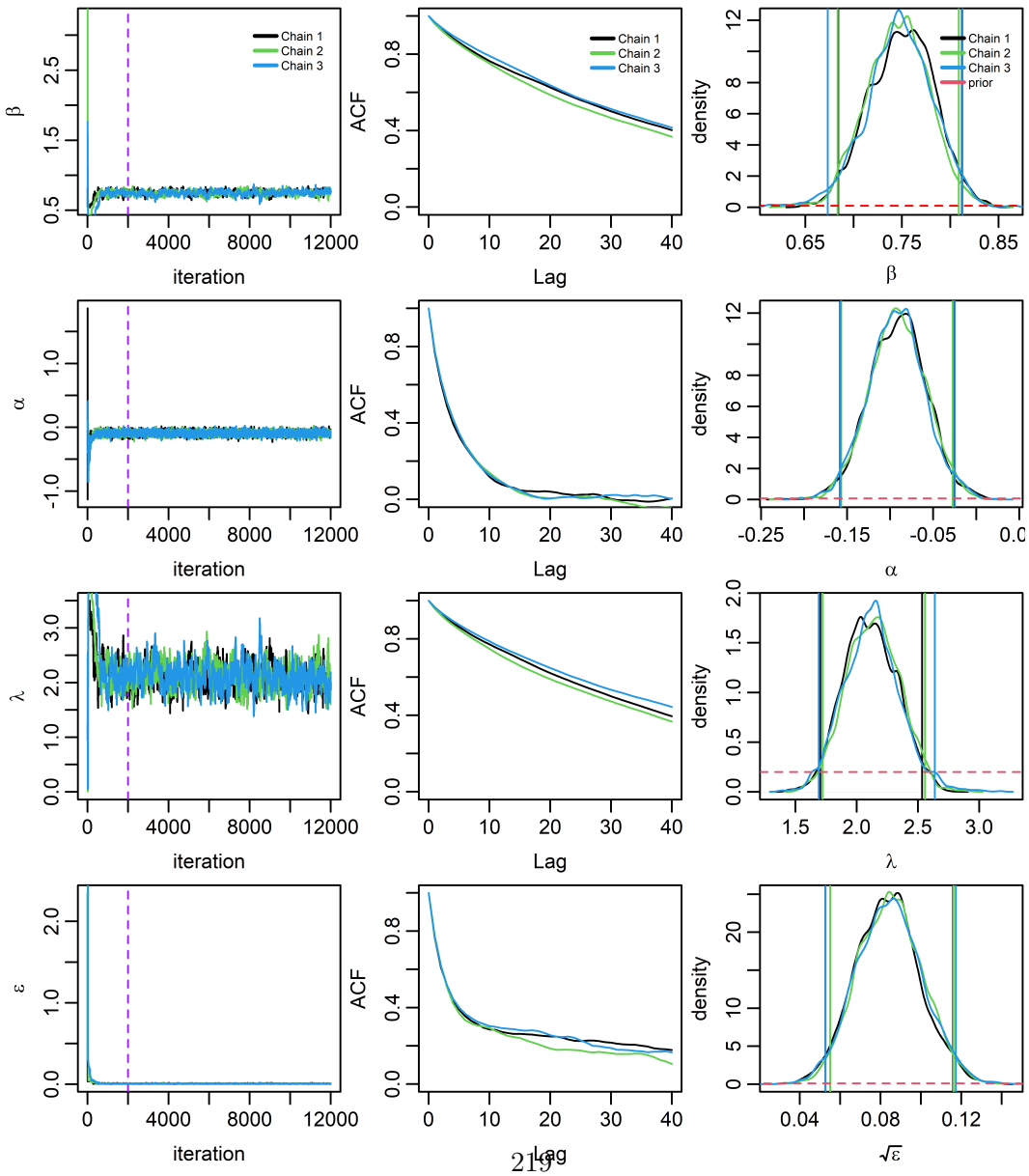


Figure B.1: Trace plots with a purple dashed line at 2000 (burn-in length), ACF curves, and uniform prior and posterior density plots with vertical lines to represent 95% Credible Intervals (CIs) for three chains of the parameters $(\beta, \alpha, \lambda, \varepsilon)$ of the model 5 ($\exp^{A}_{\alpha_1} K_{\exp}$).

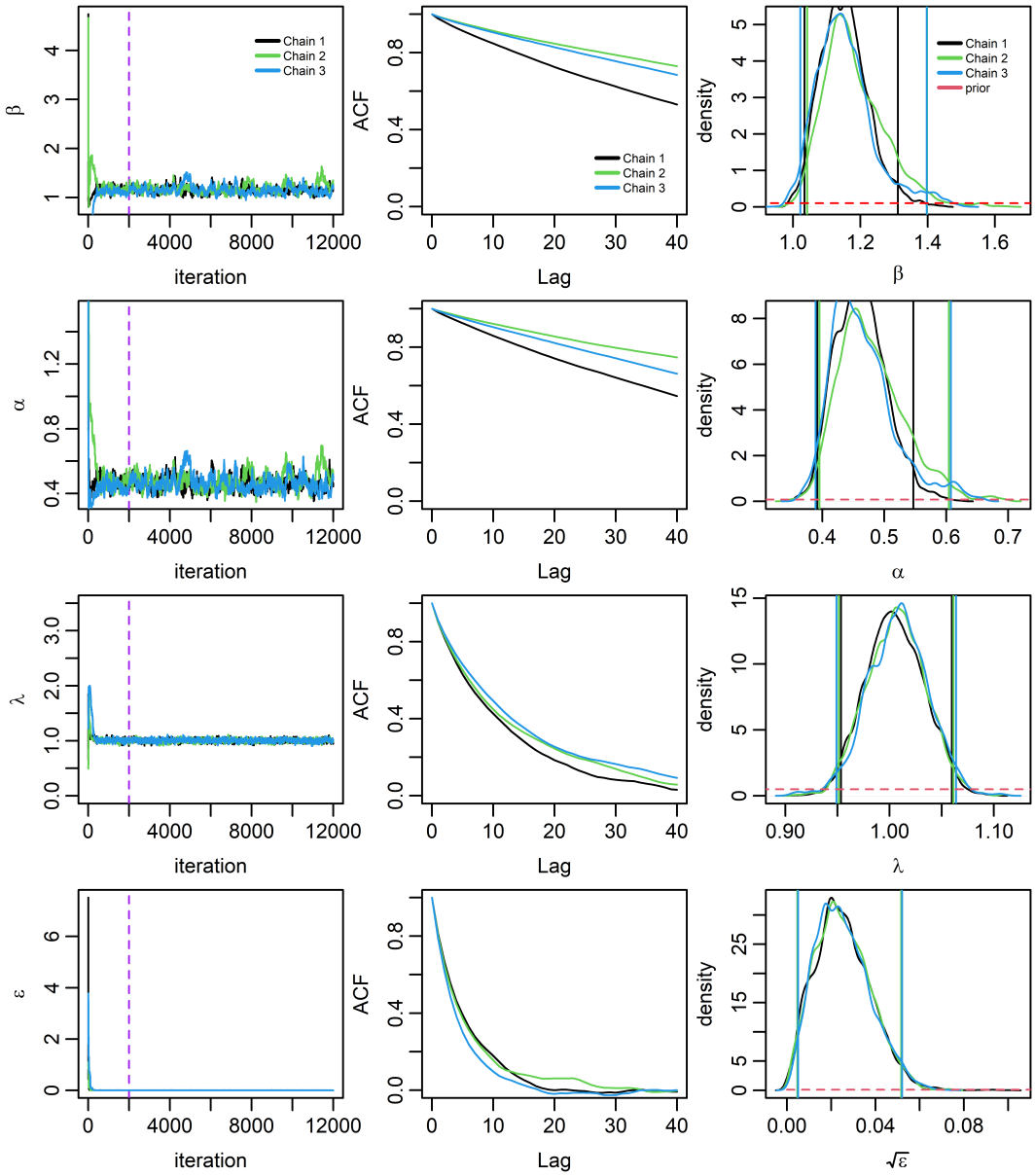


Figure B.2: Trace plots with a purple dashed line at 2000 (burn-in length), ACF curves, and uniform prior and posterior density plots with vertical lines to represent 95% Credible Intervals (CIs) for three chains of the parameters $(\beta, \alpha, \lambda, \varepsilon)$ of the model 9 ($N^{B}_{0,\alpha_2} K_{\exp}$).

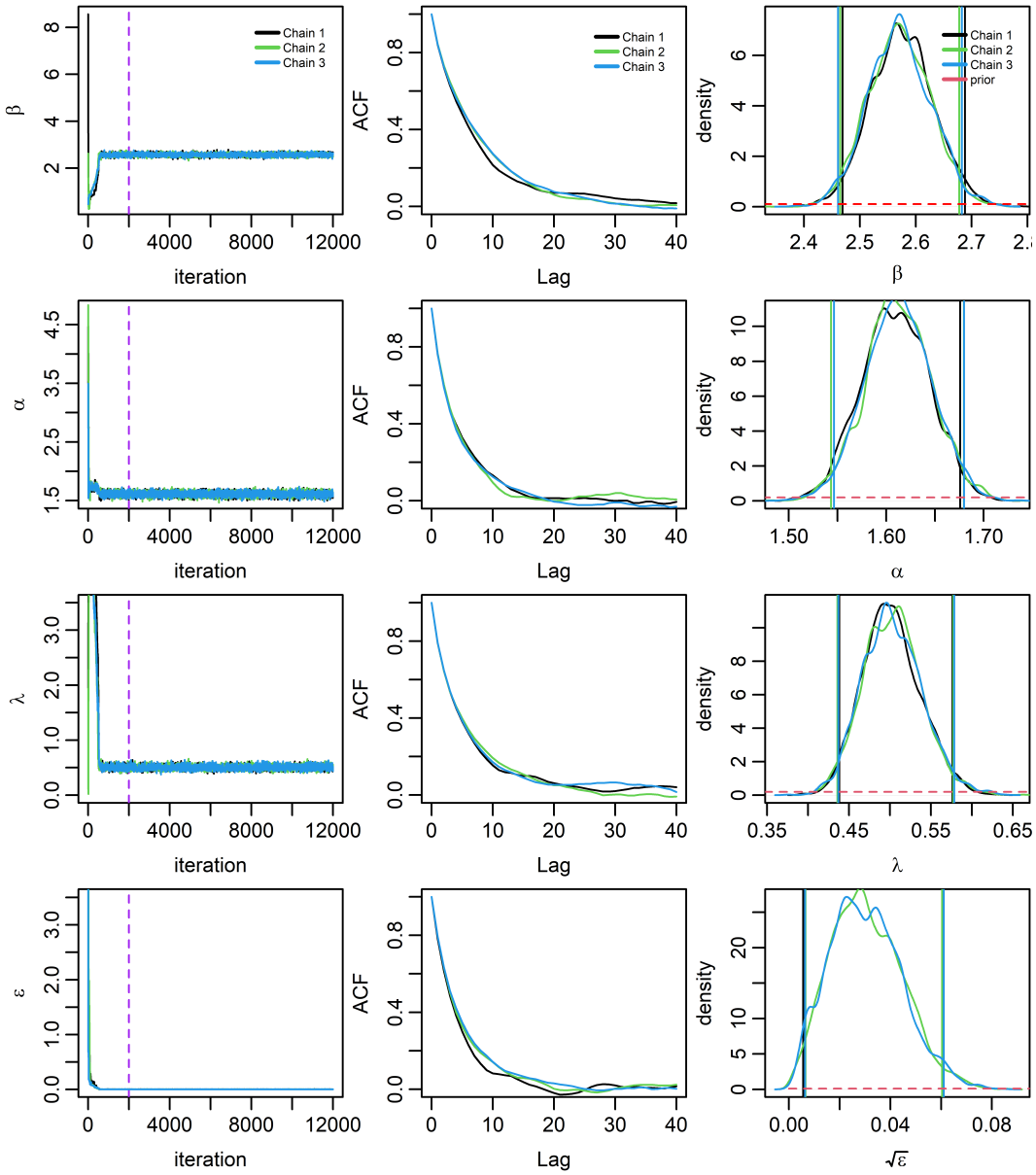


Figure B.3: Trace plots with a purple dashed line at 2000 (burn-in length), ACF curves, and uniform prior and posterior density plots with vertical lines to represent 95% Credible Intervals (CIs) for three chains of the parameters $(\beta, \alpha, \lambda, \varepsilon)$ of the model 16 $(N^{D}_{\alpha_1,.86}K_{\exp})$.

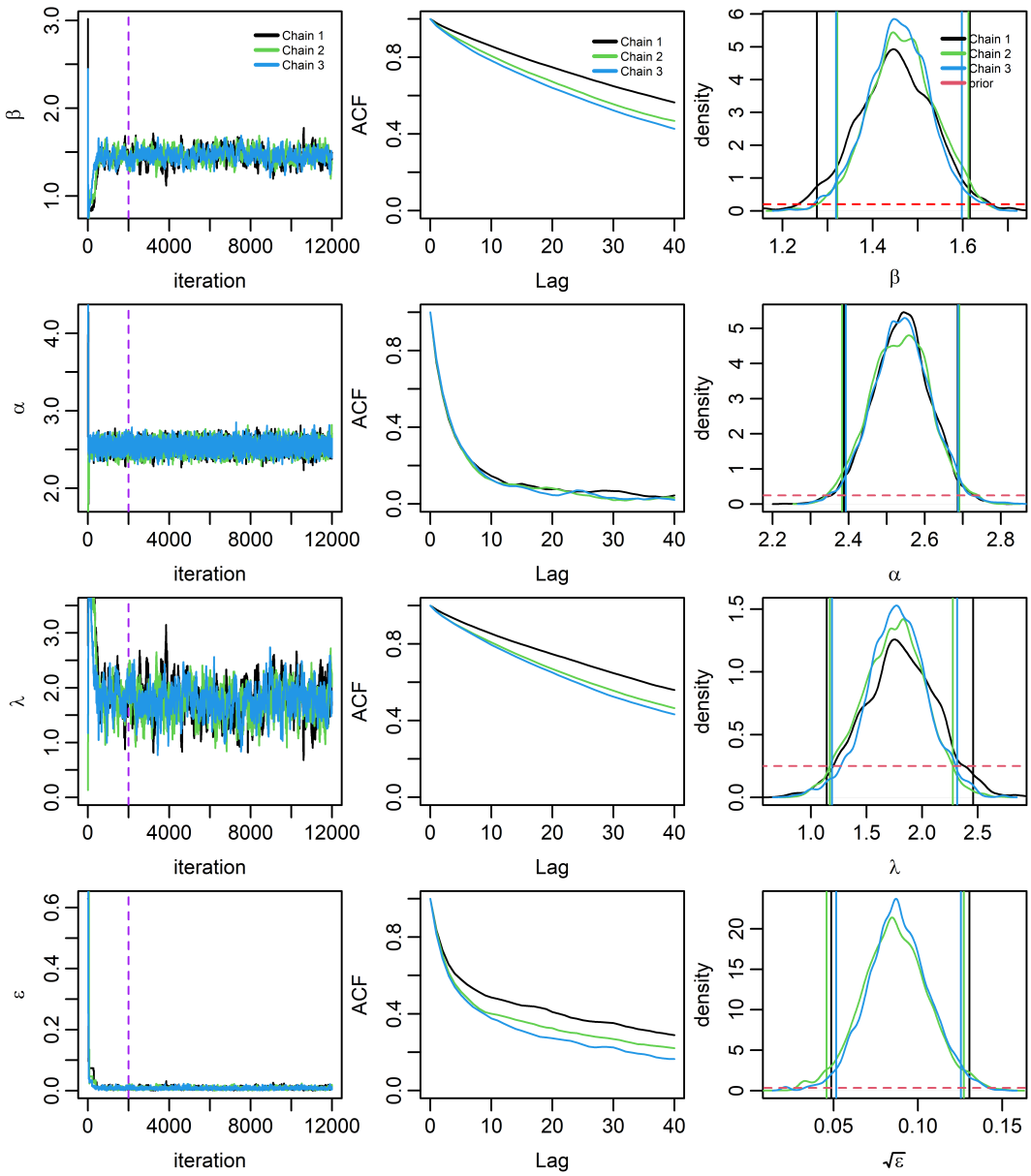


Figure B.4: Trace plots with a purple dashed line at 2000 (burn-in length), ACF curves, and uniform prior and posterior density plots with vertical lines to represent 95% Credible Intervals (CIs) for three chains of the parameters $(\beta, \alpha, \lambda, \varepsilon)$ of the model 15 $(G^{E}_{\alpha_1,.5}K_{\exp})$.

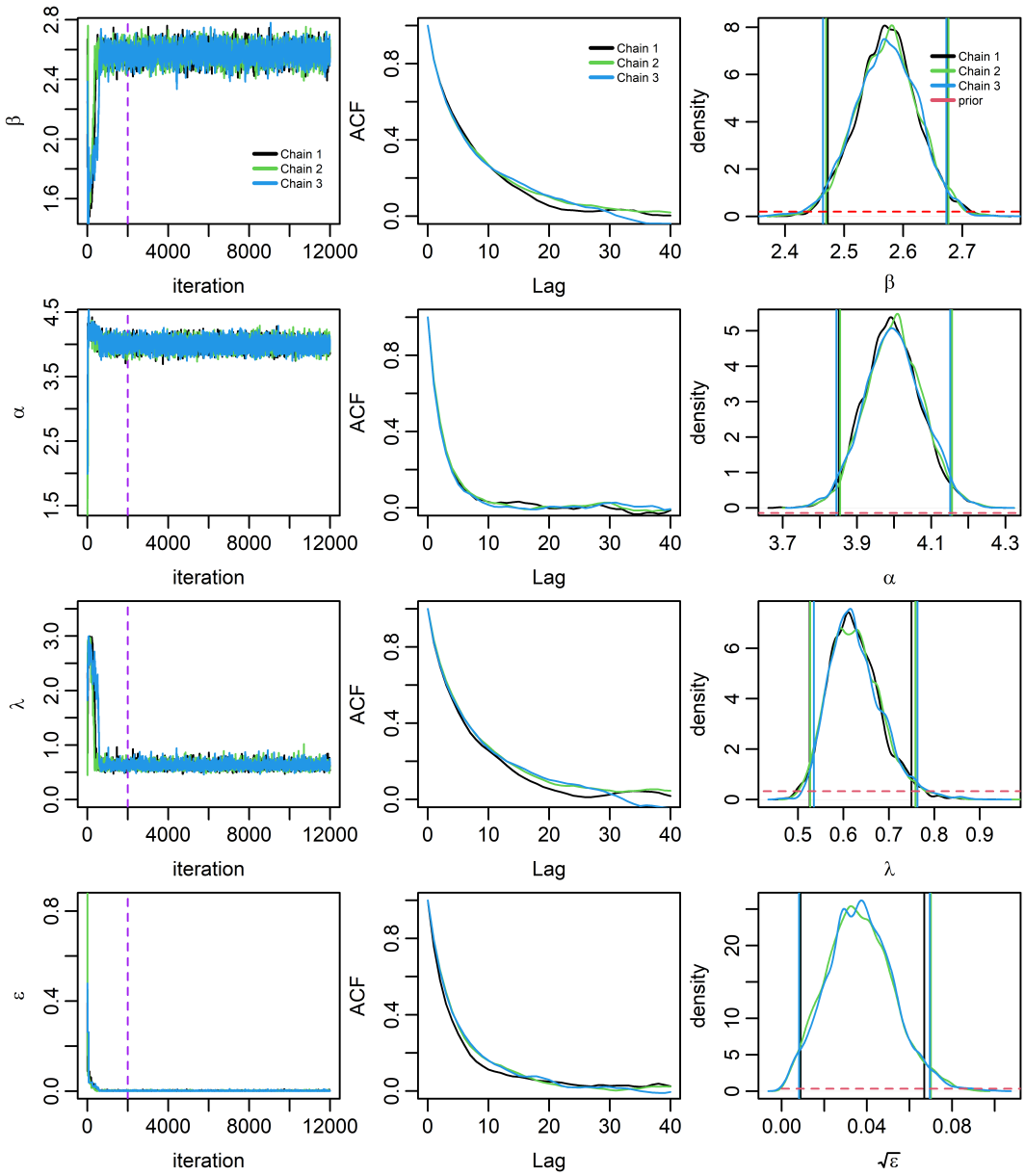


Figure B.5: Trace plots with a purple dashed line at 2000 (burn-in length), ACF curves, and uniform prior and posterior density plots with vertical lines to represent 95% Credible Intervals (CIs) for three chains of the parameters $(\beta, \alpha, \lambda, \varepsilon)$ of the model 17 $(G^D_{\alpha_1} K_{\exp})$.

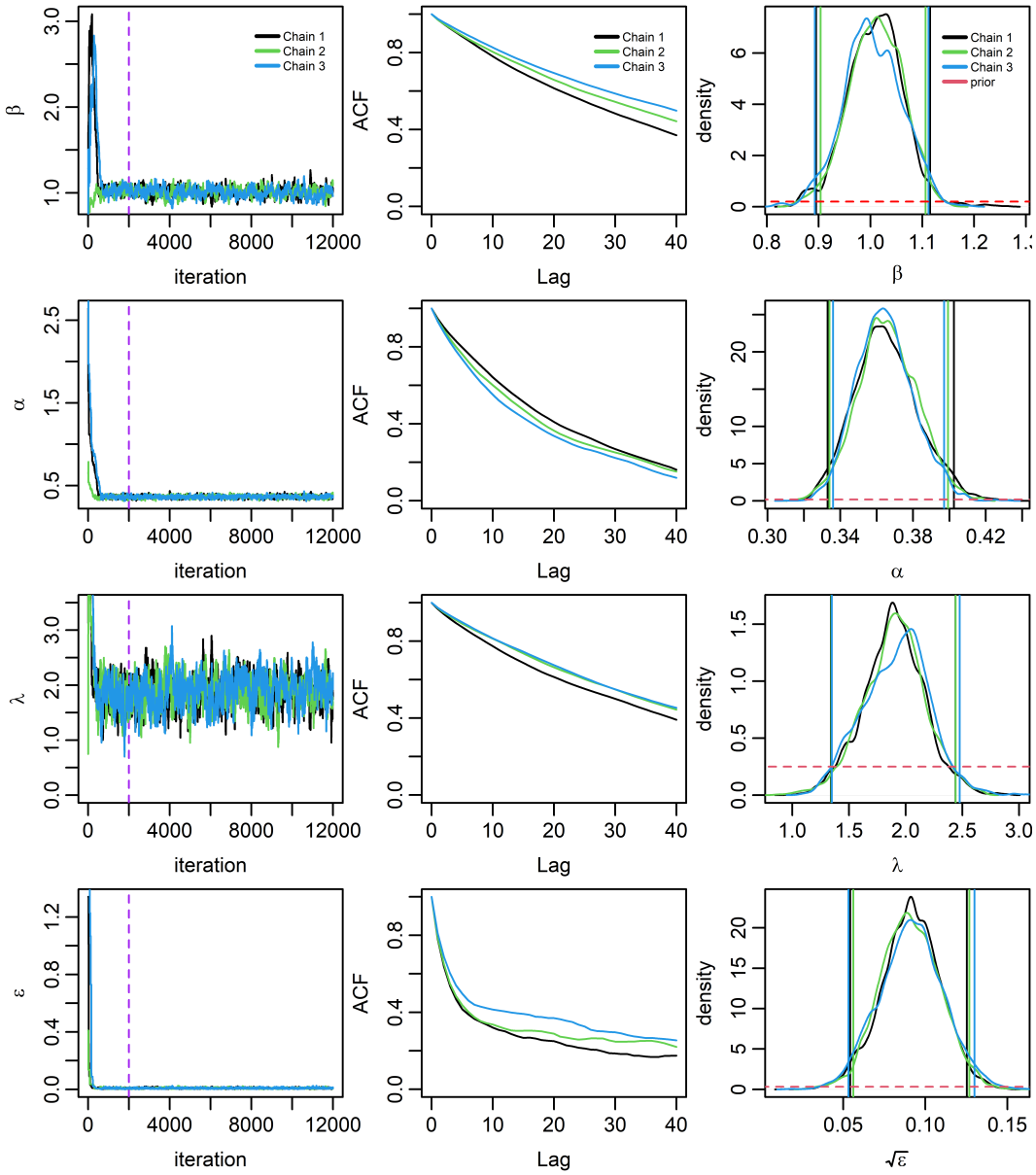


Figure B.6: Trace plots with a purple dashed line at 2000 (burn-in length), ACF curves, and uniform prior and posterior density plots with vertical lines to represent 95% Credible Intervals (CIs) for three chains of the parameters $(\beta, \alpha, \lambda, \varepsilon)$ of the model 21 $(G^F_{\alpha_1} K_{\exp})$.

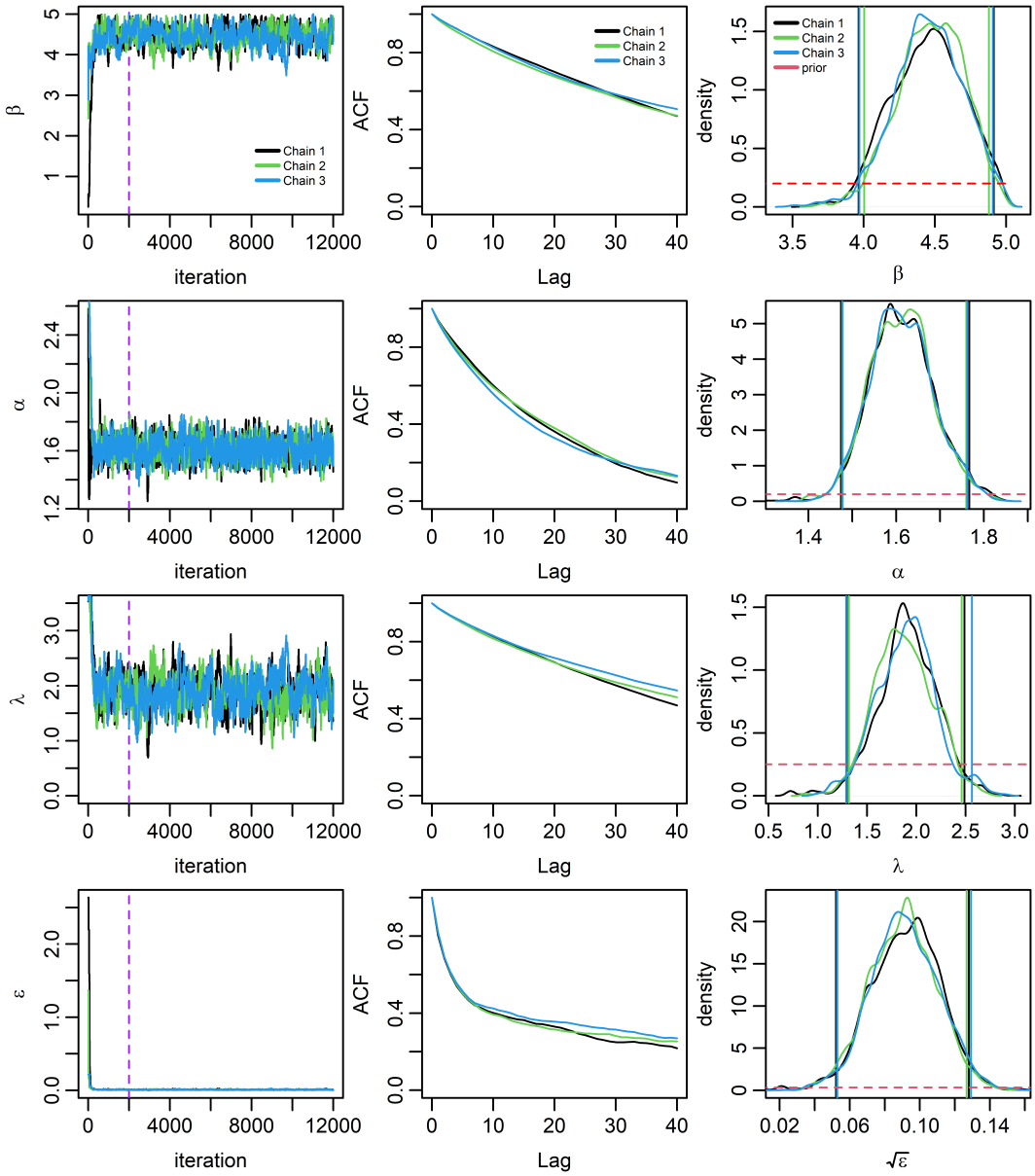


Figure B.7: Trace plots with a purple dashed line at 2000 (burn-in length), ACF curves, and uniform prior and posterior density plots with vertical lines to represent 95% Credible Intervals (CIs) for three chains of the parameters $(\beta, \alpha, \lambda, \varepsilon)$ of the model 22 ($G^{C}_{\alpha_1} K_{\exp}$).

B.2 ROC analysis

Year	J	AUC	Se	Sp	TP	FP	TN	FN
Model 1 ($\exp^{A}_{\alpha_1} K_{pwr}$)								
2003	0.85	0.96	1.00	0.71	15	246	2817	1
2004	0.9	0.98	1.00	0.94	28	305	2730	0
2005	0.95	0.99	1.00	0.93	22	164	2849	0
2006	.85	0.97	1.00	0.91	29	343	2640	1
2007	0.85	0.97	1.00	0.88	23	335	2624	1
2008	0.66	0.89	1.00	0.75	24	1009	1926	0
2009	0.84	0.96	0.99	0.89	45	343	2545	2
2010	0.75	0.91	1.00	0.76	45	662	2180	1
2011	0.73	0.92	1.00	0.74	60	750	2032	0
2012	0.68	0.89	0.99	0.69	68	748	1961	5
2013	0.65	0.89	0.97	0.82	78	474	2141	16
2014	0.58	0.85	0.98	0.73	118	752	1792	16
2015	0.63	0.86	0.99	0.72	121	696	1654	10
2016	0.61	0.88	0.96	0.67	104	493	1732	21
2017	0.63	0.87	0.99	0.77	90	596	1530	9
2018	0.56	0.85	0.99	0.67	83	690	1344	9
2019	0.67	0.89	0.97	0.79	99	402	1519	14

Table B.1: Evaluation the model performance of the model 1 ($\exp^{A}_{\alpha_1} K_{pwr}$) using different performance metrics: Youdon index,(J), Area Under ROC Curve (AUC), sensitivity (Se), specificity (Sp), number of true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN).

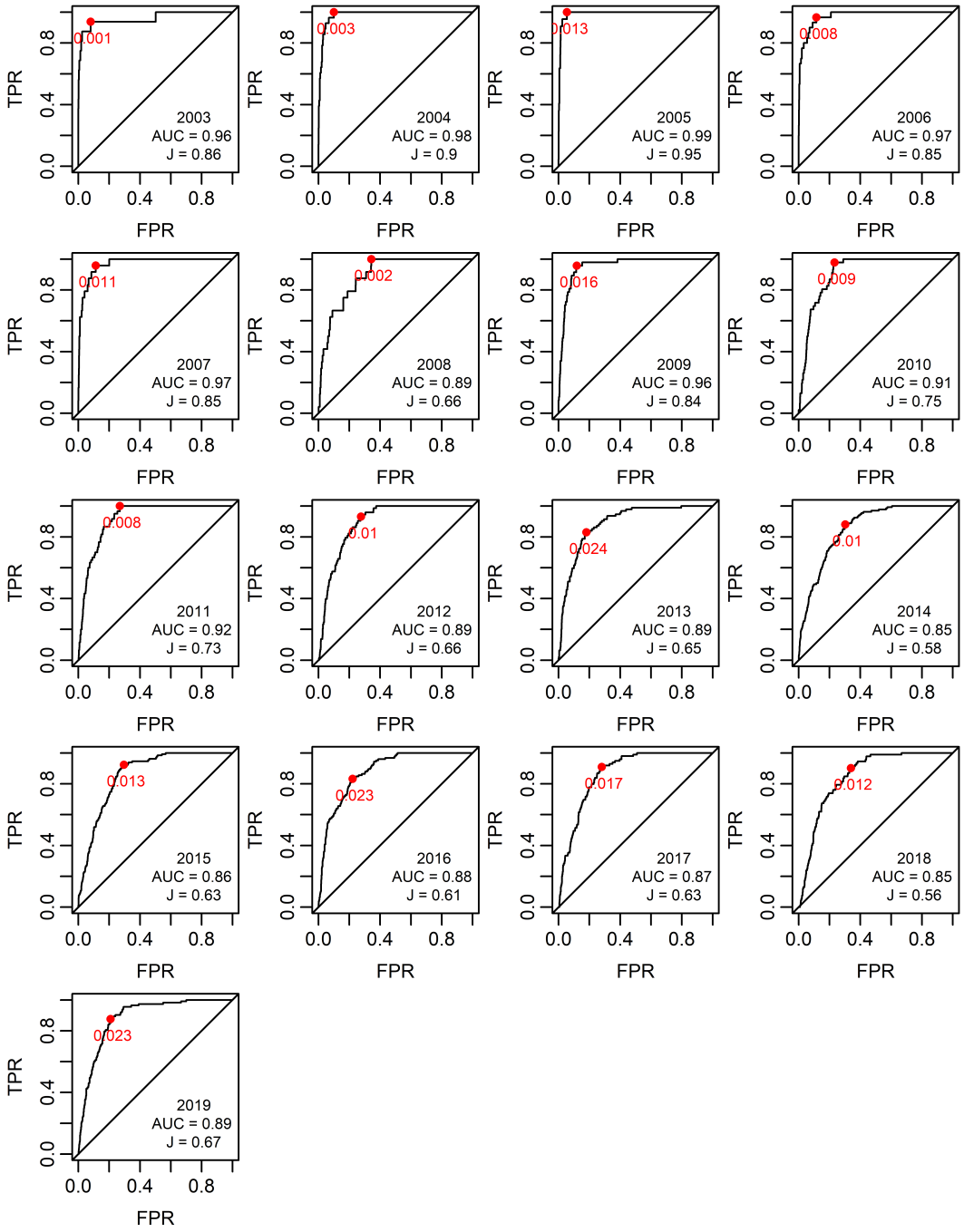


Figure B.8: Evaluation the model performance of the model 1 ($\exp^{A}_{\alpha_1} K_{pwr}$) using different performance metrics: Youdon index,(J), Area Under ROC Curve (AUC), sensitivity (Se), specificity (Sp), number of true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN). The sensitivity is the TPR, while the specificity is 1-FPR.

Year	J	AUC	Se	Sp	TP	FP	TN	FN
Model 5 ($\exp^{A}_{\alpha_1} K_{\exp}$)								
2003	0.85	0.93	0.875	0.978778975	14	65	2998	2
2004	0.92	0.98	1	0.924876442	28	228	2807	0
2005	0.97	0.99	1	0.973116495	22	81	2932	0
2006	0.90	0.98	0.966666667	0.937646664	29	186	2797	1
2007	.86	0.97	1	0.861101724	24	411	2548	0
2008	0.71	0.91	0.958333333	0.750255537	23	733	2202	1
2009	0.84	0.95	0.936170213	0.902700831	44	281	2607	3
2010	0.77	0.92	0.97826087	0.791695989	45	592	2250	1
2011	0.77	0.93	1	0.785046729	60	598	2184	0
2012	0.72	0.90	0.95890411	0.756736803	70	659	2050	3
2013	0.67	0.90	0.85106383	0.817208413	80	478	2137	14
2014	0.59	0.86	0.902985075	0.682789198	121	787	1694	13
2015	0.68	0.88	0.93129771	0.748510638	122	591	1759	9
2016	0.64	0.90	0.88	0.765842697	110	521	1704	15
2017	0.70	0.89	0.929292929	0.765757291	92	498	1628	7
2018	0.66	0.86	0.956521739	0.700589971	88	609	1425	4
2019	0.73	0.90	0.946902655	0.78292556	107	417	1504	6
			0.931778929	0.827962012	1079	7735	37226	79

Table B.2: Evaluation of Model Accuracy using Different Performance Metrics: Youdon index,(J), Area Under ROC Curve (AUC), Sensitivity (Sen.), Specificity (Spe.), number of True Positives (TP), False Positives (FP), True Negatives (TN), and False Negatives (FN). By comparing J, AUC, FP, and FN between models, the bold value indicates better performance or no changes between different models.

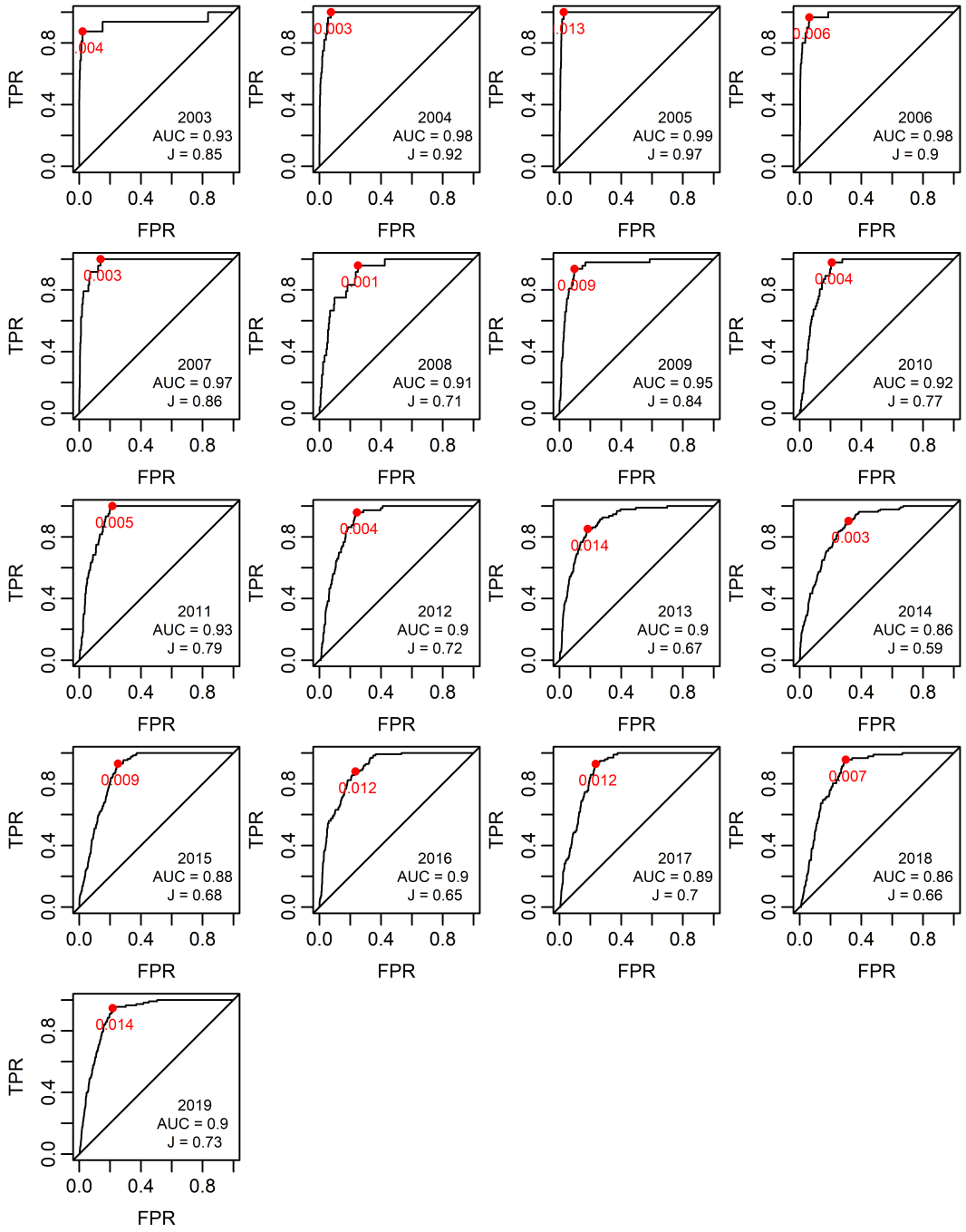


Figure B.9: Evaluation the model performance of the model 5 ($\exp^{A}_{\alpha_1} K_{\exp}$) using different performance metrics: Youdon index,(J), Area Under ROC Curve (AUC), sensitivity (Se), specificity (Sp), number of true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN). The sensitivity is the TPR, while the specificity is 1-FPR.

Year	J	AUC	Se	Sp	TP	FP	TN	FN
Model 16 ($N^D_{\alpha_1,.86}K_{\exp}$)								
2003	.85	.96	0.94	0.92	15	254	2809	1
2004	0.88	0.96	0.96	0.90	27	307	2728	1
2005	0.90	0.97	1.00	0.90	22	304	2709	0
2006	0.85	0.96	0.97	0.89	29	343	2640	1
2007	0.82	0.95	0.92	0.91	22	277	2682	2
2008	0.84	0.88	0.88	0.77	21	684	2251	3
2009	0.84	0.95	0.96	0.88	45	338	2550	2
2010	0.71	0.90	0.98	0.73	45	771	2071	1
2011	0.70	0.90	0.98	0.71	59	801	1981	1
2012	0.67	0.88	0.96	0.71	70	795	1914	3
2013	0.65	0.88	0.85	0.80	80	522	2093	14
2014	0.58	0.85	0.90	0.68	121	791	1690	13
2015	0.62	0.85	0.91	0.71	119	686	1664	12
2016	0.82	0.88	0.94	0.68	118	712	1513	7
2017	0.83	0.88	0.90	0.73	89	568	1558	10
2018	0.80	0.85	0.91	0.69	84	636	1398	8
2019	0.71	089	0.96	0.75	108	474	1447	5

Table B.3: Evaluation the model performance of the model 16 ($N^D_{\alpha_1,.86}K_{\exp}$) using different performance metrics: Youdon index,(J), Area Under ROC Curve (AUC), sensitivity (Se), specificity (Sp), number of true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN).

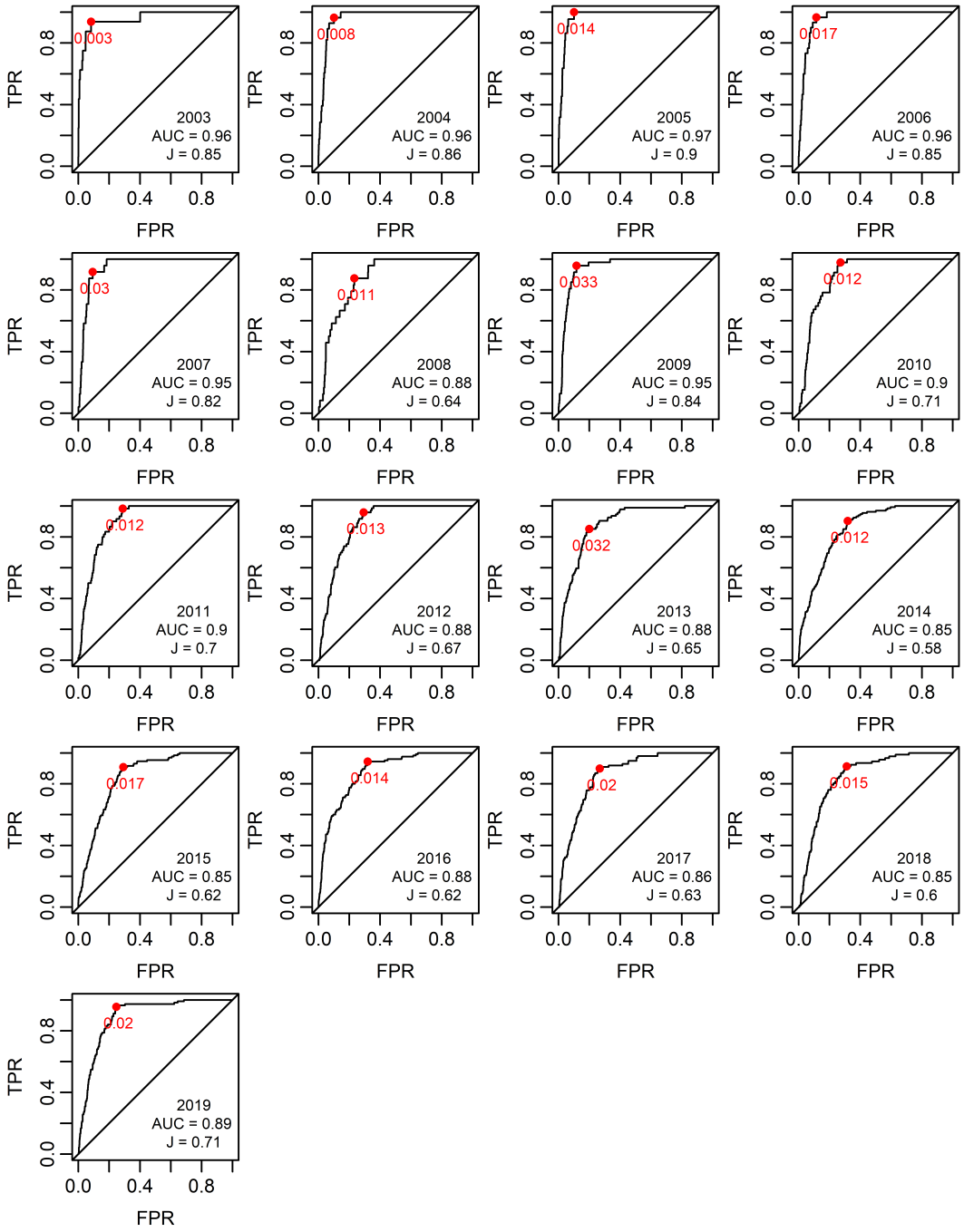


Figure B.10: Evaluation the model performance of the model 16 ($N^{D}_{\alpha_1,.86}K_{\exp}$) using different performance metrics: Youdon index,(J), Area Under ROC Curve (AUC), sensitivity (Se), specificity (Sp), number of true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN). The sensitivity is the TPR, while the specificity is 1-FPR.

Year	J	AUC	Se	Sp	TP	FP	TN	FN
Model 15 ($G^E_{\alpha_1,.5}K_{\exp}$)								
2003	0.85	0.94	0.88	0.98	14	69	2994	2
2004	0.92	0.98	1	0.92	28	233	2802	0
2005	0.97	0.99	1	0.97	22	84	2929	0
2006	0.90	0.98	0.97	0.94	29	190	2793	1
2007	0.88	0.97	1	0.88	24	355	2604	0
2008	0.67	0.90	0.96	0.71	23	848	2087	1
2009	0.83	0.95	0.94	0.90	44	296	2592	3
2010	0.76	0.92	0.96	0.80	44	561	2281	2
2011	0.77	0.92	0.97	0.80	58	550	2232	2
2012	0.71	0.90	0.96	0.75	70	675	2034	3
2013	0.65	0.89	0.84	0.81	79	485	2130	15
2014	0.59	0.86	0.95	0.64	127	887	1594	7
2015	0.66	0.87	0.92	0.74	121	621	1729	10
2016	0.64	0.89	0.94	0.70	117	658	1567	8
2017	0.67	0.89	0.91	0.76	90	511	1615	9
2018	0.63	0.86	0.90	0.73	83	550	1484	9
2019	0.73	0.91	0.96	0.77	108	433	1488	5

Table B.4: Evaluation the model performance of the model 15 ($G^E_{\alpha_1,.5}K_{\exp}$) using different performance metrics: Youdon index,(J), Area Under ROC Curve (AUC), sensitivity (Se), specificity (Sp), number of true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN).

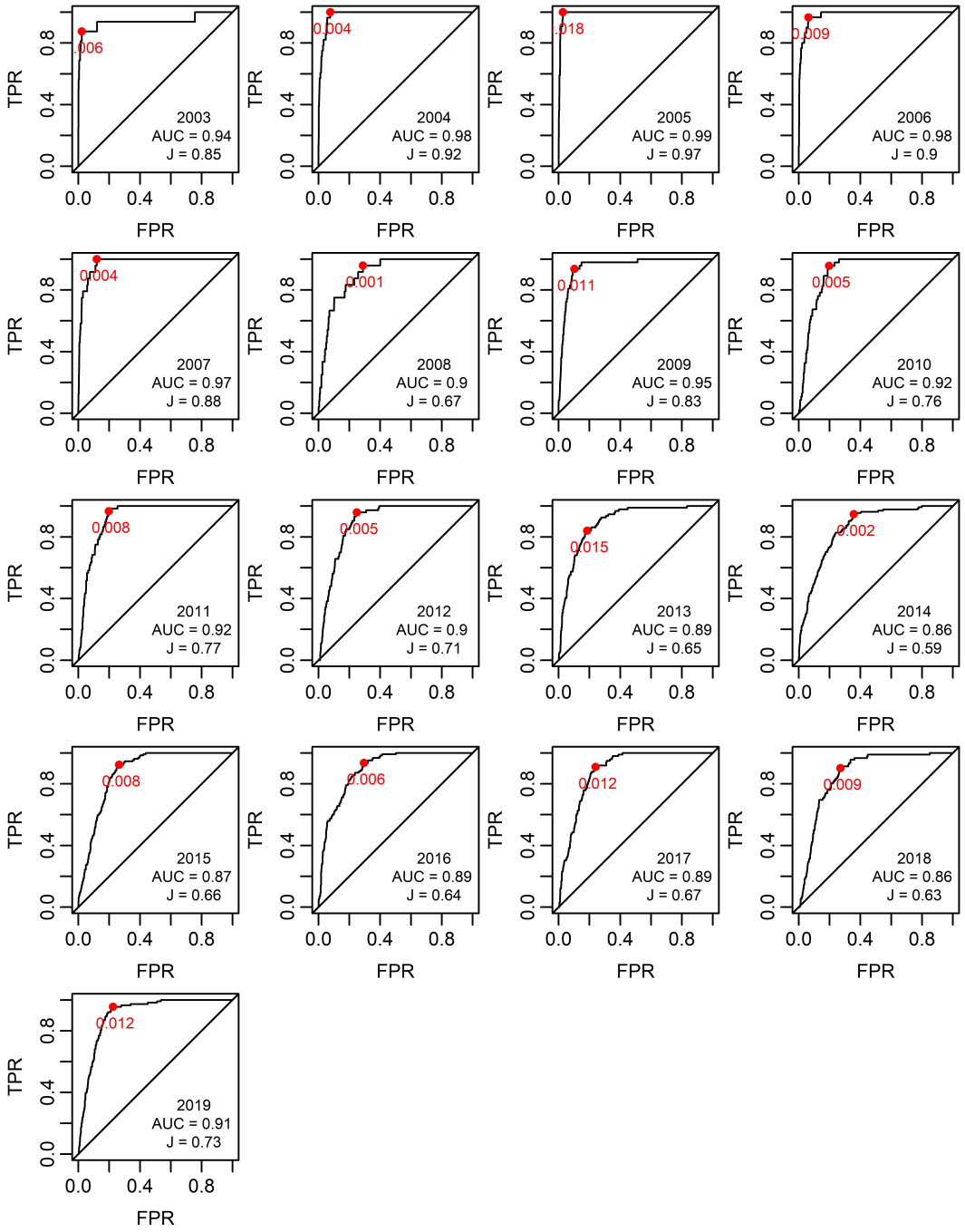


Figure B.11: Evaluation the model performance of the model 15 ($G^{E}_{\alpha_1,.5}K_{\exp}$) using different performance metrics: Youdon index,(J), Area Under ROC Curve (AUC), sensitivity (Se), specificity (Sp), number of true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN). The sensitivity is the TPR, while the specificity is 1-FPR.

Year	J	AUC	Sen.	Spe.	TP	FP	TN	FN
Model 17 ($G^D_{\alpha_1} K_{\exp}$)								
2003	0.86	0.96	0.94	0.92	15	231	2832	1
2004	0.88	0.97	1.00	0.88	28	351	2684	0
2005	0.92	0.96	1.00	0.92	22	256	2757	0
2006	0.87	0.97	0.97	0.90	29	296	2687	1
2007	0.84	0.96	1.00	0.84	24	487	2472	0
2008	0.66	0.88	0.88	0.78	21	636	2299	3
2009	0.84	0.95	0.96	0.89	45	327	2561	2
2010	0.72	0.90	0.98	0.74	45	736	2106	1
2011	0.71	0.91	0.98	0.72	59	770	2012	1
2012	0.67	0.88	0.96	0.72	70	770	1939	3
2013	0.65	0.88	0.84	0.81	4 79	500	2115	15
2014	0.58	0.85	0.93	0.65	124	856	1625	10
2015	0.61	0.85	0.92	0.70	120	712	1638	11
2016	0.60	0.87	0.94	0.66	118	766	1459	7
2017	0.62	0.86	0.90	0.72	89	592	1534	10
2018	0.58	0.85	0.89	0.69	82	634	1400	10
2019	0.69	0.89	0.96	0.73	108	514	1407	5

Table B.5: Evaluation the model performance of the model 17 ($G^D_{\alpha_1} K_{\exp}$) using different performance metrics: Youdon index,(J), Area Under ROC Curve (AUC), sensitivity (Se), specificity (Sp), number of true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN).

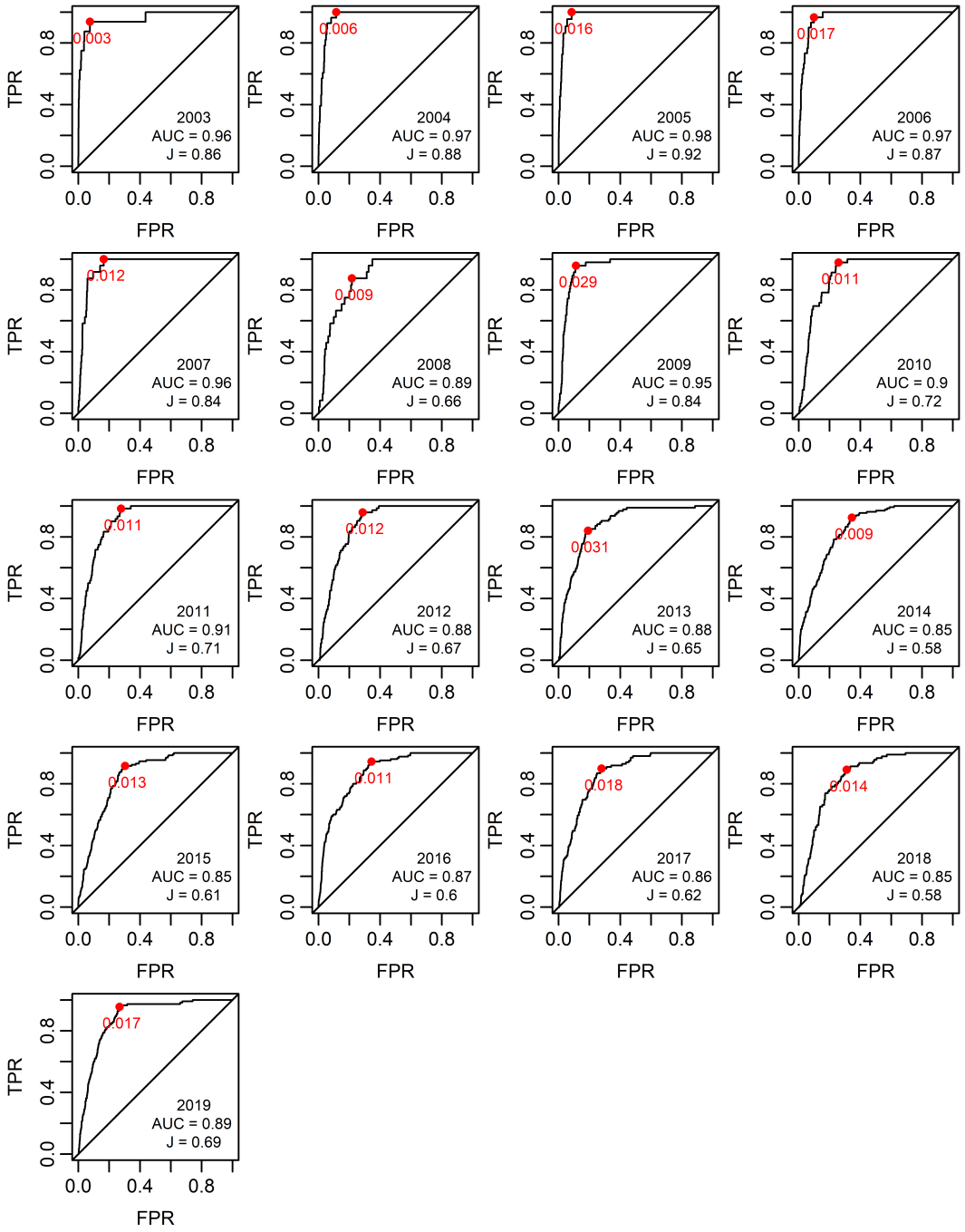


Figure B.12: Evaluation the model performance of the model 17 ($G^D_{\alpha_1} K_{\exp}$) using different performance metrics: Youdon index,(J), Area Under ROC Curve (AUC), sensitivity (Se), specificity (Sp), number of true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN). The sensitivity is the TPR, while the specificity is 1-FPR.

Year	J	AUC	Sen.	Spe.	TP	FP	TN	FN
Model 21 ($G^F_{\alpha_1} K_{\exp}$)								
2003	0.78	0.92	0.88	0.90	14	291	2772	2
2004	091	0.99	0.96	0.95	27	154	2881	1
2005	0.93	0.99	1.00	0.93	22	197	2816	0
2006	0.83	0.97	0.93	0.90	28	307	2676	2
2007	0.81	0.97	0.96	0.86	23	425	2534	1
2008	0.64	0.86	0.79	0.84	19	460	2475	5
2009	0.83	0.95	0.96	0.87	45	380	2508	2
2010	0.71	0.91	0.96	0.76	44	692	2150	2
2011	0.74	0.91	0.98	0.76	59	670	2112	1
2012	0.66	0.88	0.90	0.75	66	671	2038	7
2013	0.66	0.89	0.91	0.74	86	674	1941	8
2014	0.59	0.83	0.81	0.78	109	547	1934	25
2015	0.61	0.85	0.87	0.74	114	609	1741	17
2016	0.62	0.86	0.88	0.74	110	588	1637	15
2017	0.65	0.87	0.87	0.78	86	470	1656	13
2018	0.63	0.85	0.85	0.78	78	445	1589	14
2019	0.73	0.91	0.98	0.75	111	477	1444	2

Table B.6: Evaluation the model performance of the model 21 ($G^F_{\alpha_1} K_{\exp}$) using different performance metrics: Youdon index,(J), Area Under ROC Curve (AUC), sensitivity (Se), specificity (Sp), number of true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN).

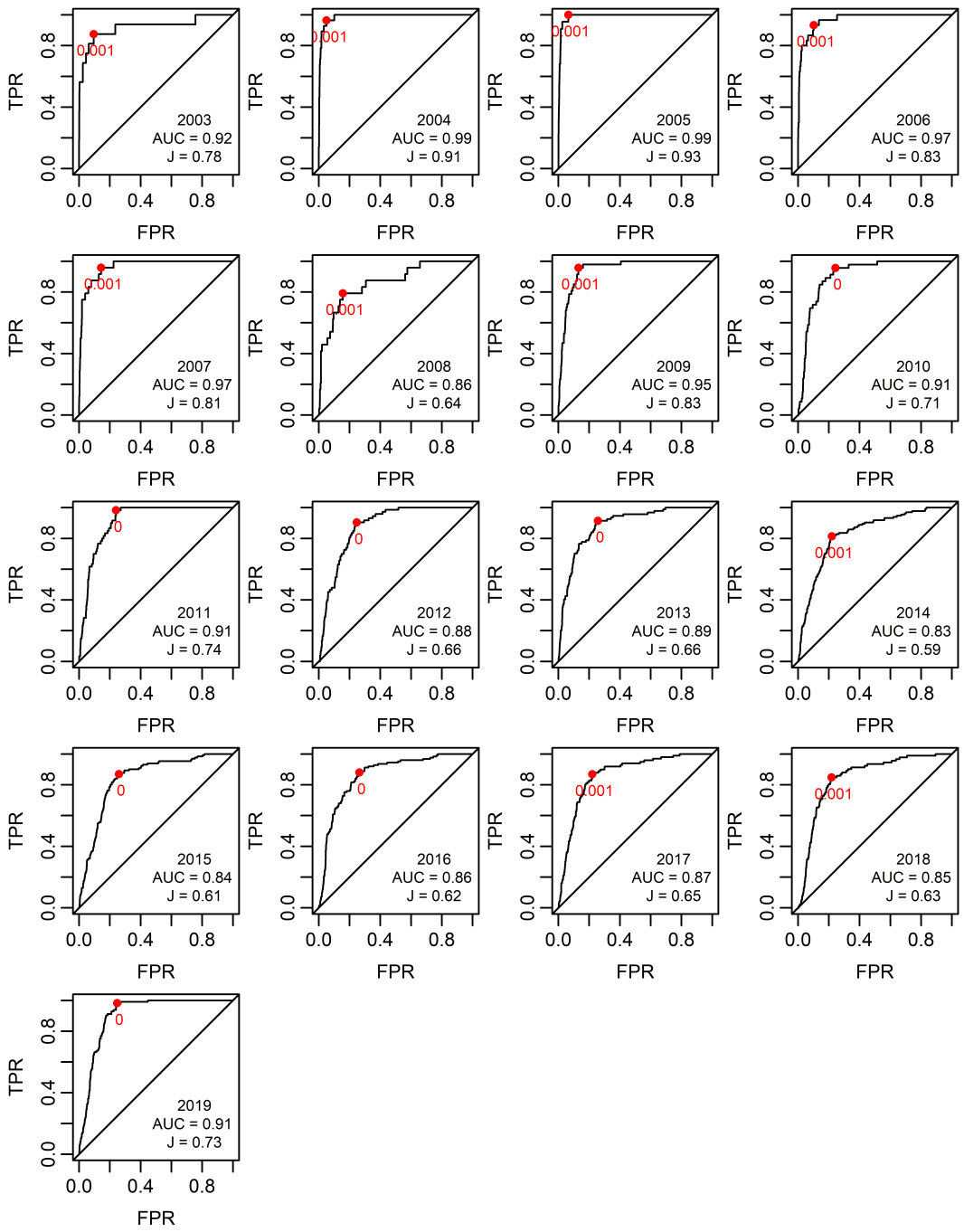


Figure B.13: Evaluation the model performance of the model 21 ($G^F_{\alpha_1} K_{\exp}$) using different performance metrics: Youdon index,(J), Area Under ROC Curve (AUC), sensitivity (Se), specificity (Sp), number of true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN). The sensitivity is the TPR, while the specificity is 1-FPR.

Year	J	AUC	Se	Sp	TP	FP	TN	FN
Model 22 ($G^C_{\alpha_1} K_{\exp}$)								
2003	0.72	0.90	0.81	0.91	13	283	2780	3
2004	0.93	0.99	1.00	0.93	28	201	2834	0
2005	0.87	0.98	0.95	0.91	21	265	2748	1
2006	.75	0.95	0.87	0.89	26	342	2641	4
2007	0.77	0.94	0.96	0.81	23	569	2390	1
2008	0.54	0.83	0.79	0.75	19	745	2190	5
2009	0.66	0.90	0.87	0.79	41	607	2281	6
2010	0.70	0.88	0.91	0.79	42	595	2247	4
2011	0.57	0.86	0.70	0.87	42	355	2427	18
2012	0.46	0.80	0.84	0.62	61	1020	1689	12
2013	0.56	0.84	0.74	0.81	70	491	2124	24
2014	0.44	0.77	0.72	0.71	97	715	1766	37
2015	0.38	0.75	0.84	0.54	110	1092	1258	21
2016	0.42	0.78	0.78	0.64	98	812	1413	27
2017	0.48	0.79	0.77	0.71	76	616	1510	23
2018	0.42	0.76	0.87	0.55	80	919	1115	12
2019	0.49	0.79	0.97	0.52	110	929	992	3

Table B.7: Evaluation the model performance of the model 22 ($G^C_{\alpha_1} K_{\exp}$) using different performance metrics: Youdon index,(J), Area Under ROC Curve (AUC), sensitivity (Se), specificity (Sp), number of true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN).

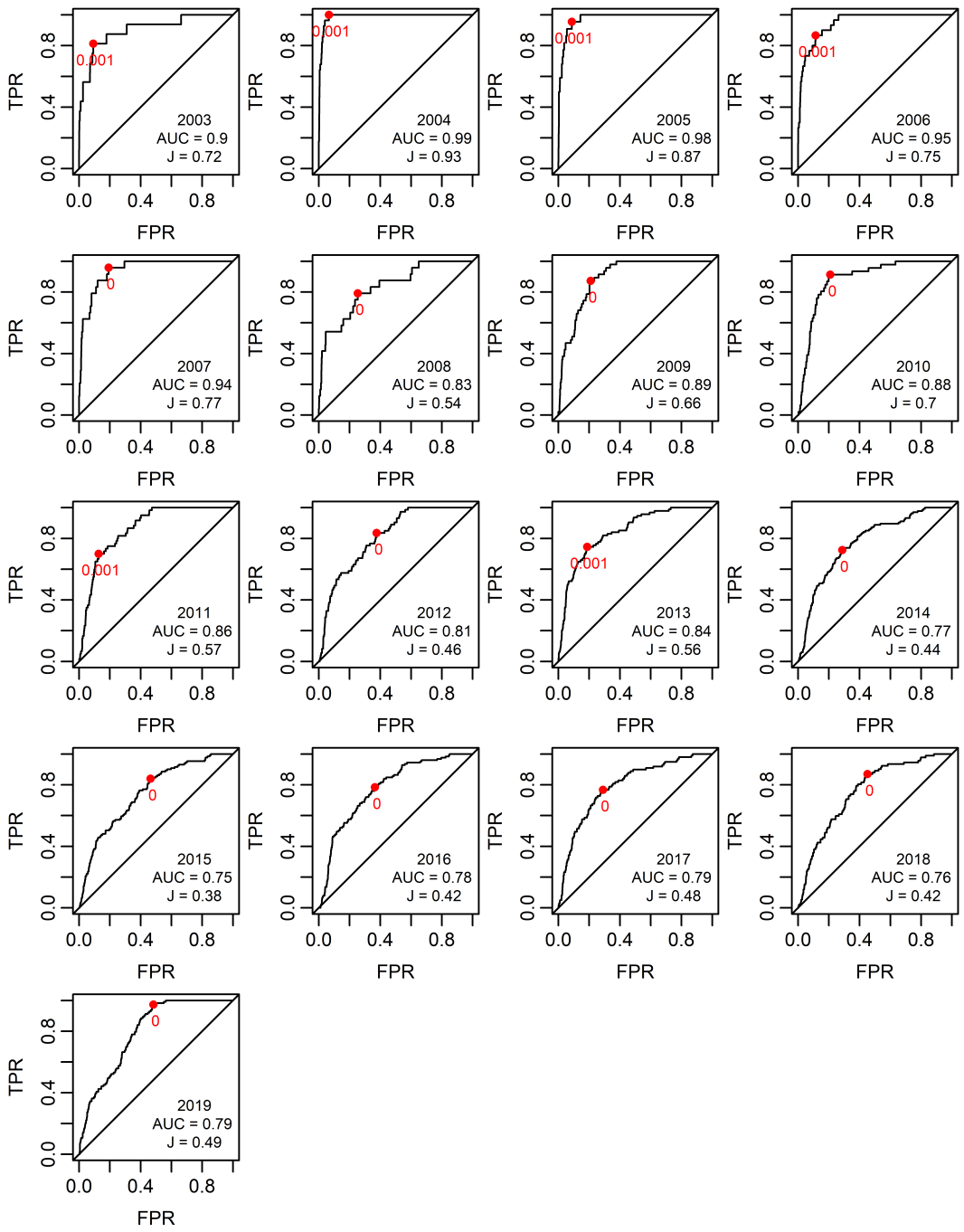


Figure B.14: Evaluation the model performance of the model 22 ($G^{C}_{\alpha_1} K_{\exp}$) using different performance metrics: Youdon index,(J), Area Under ROC Curve (AUC), sensitivity (Se), specificity (Sp), number of true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN). The sensitivity is the TPR, while the specificity is 1-FPR.

Year	J	AUC	Se	Sp	TP	FP	TN	FN
Model 24 ($G^E_{\alpha_1} K_{\exp}$)								
2003	0.82	0.93	0.88	0.95	14	157	2906	2
2004	0.93	0.99	0.96	0.97	27	105	2930	1
2005	0.97	0.99	1.00	0.97	22	105	2908	0
2006	0.87	0.98	0.97	0.90	29	292	2691	1
2007	0.83	0.97	1.00	0.83	24	514	2445	0
2008	0.65	0.88	0.79	0.86	19	424	2511	5
2009	0.84	0.95	0.98	0.86	46	392	2499	1
2010	0.75	0.92	0.96	0.79	44	594	2248	2
2011	0.78	0.92	1.00	0.78	60	620	2162	0
2012	0.68	0.89	0.96	0.72	70	767	1942	3
2013	0.66	0.89	0.91	0.74	86	672	1943	8
2014	0.60	0.85	0.81	0.78	109	538	1943	25
2015	0.64	0.86	0.89	0.76	116	569	1781	15
2016	0.64	0.88	0.91	0.72	114	616	1609	11
2017	0.66	0.88	0.91	0.75	90	529	1597	9
2018	0.63	0.86	0.91	0.71	84	583	1451	8
2019	0.74	0.91	0.96	0.77	109	441	1480	4

Table B.8: Evaluation the model performance of the model 24 ($G^E_{\alpha_1} K_{\exp}$) using different performance metrics: Youdon index,(J), Area Under ROC Curve (AUC), sensitivity (Se), specificity (Sp), number of true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN).

B.3 Suitability maps

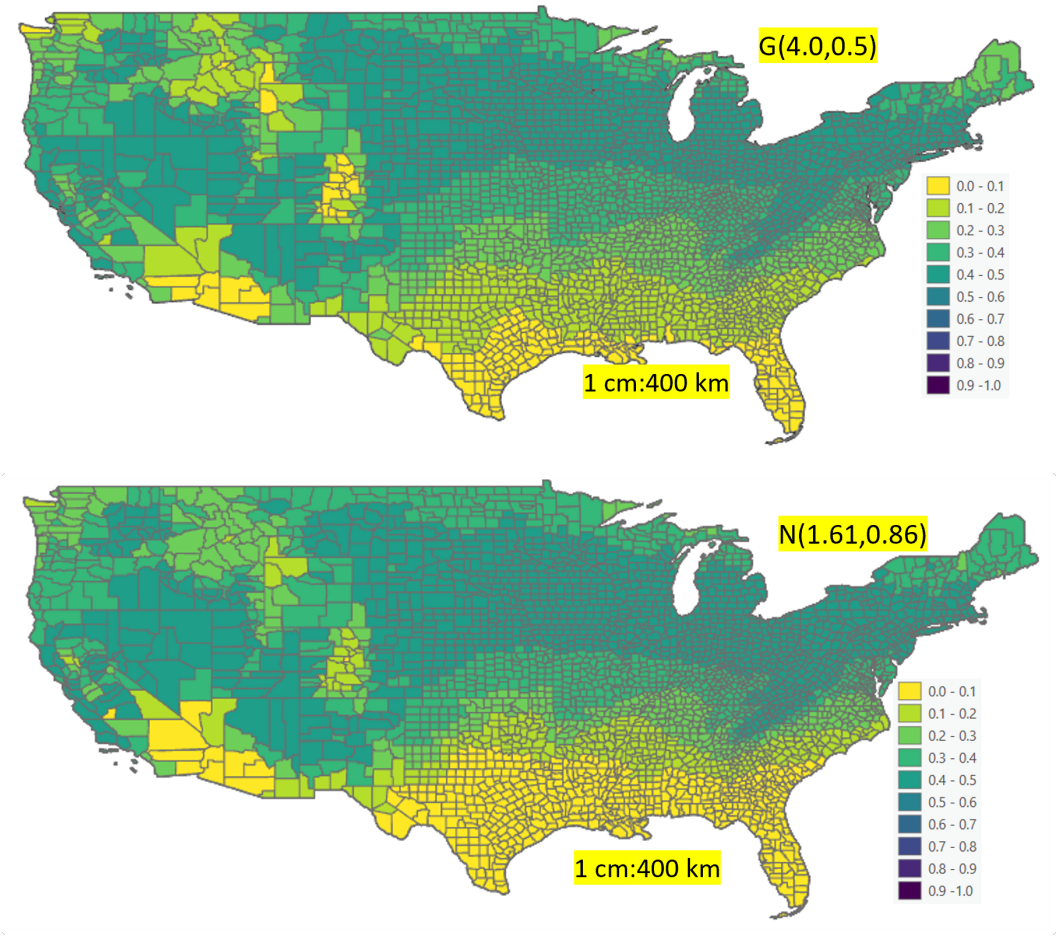


Figure B.15: The overall suitability for each county in the USA for the invasive EAB colonisation under the model 17 ($G^D_{\alpha_1} K_{\exp}$) and model 16 ($N^D_{\alpha_1,.86} K_{\exp}$), respectively. Each of these model is the spatio-temporal colonisation–dispersal model based on estimated ash density and the MAGDD. The value calculated for each county is the mean of the posterior distribution of the suitability function.

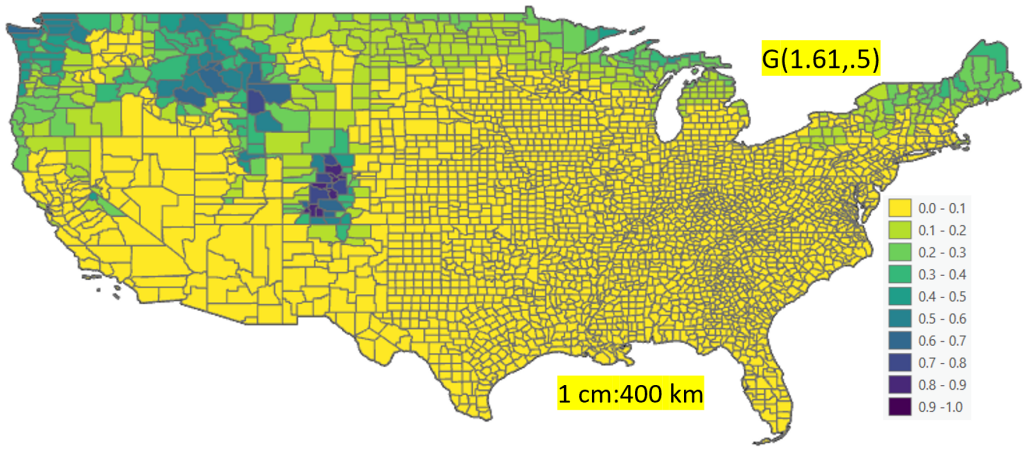


Figure B.16: The Overall suitability for each county in the USA for the invasive EAB colonisation under the model 22 ($G^{C}_{\alpha_1} K_{\exp}$) which is the spatio-temporal colonisation–dispersal model based on estimated ash density and the MAGDD. The value calculated for each county is the mean of the posterior distribution of the suitability function.

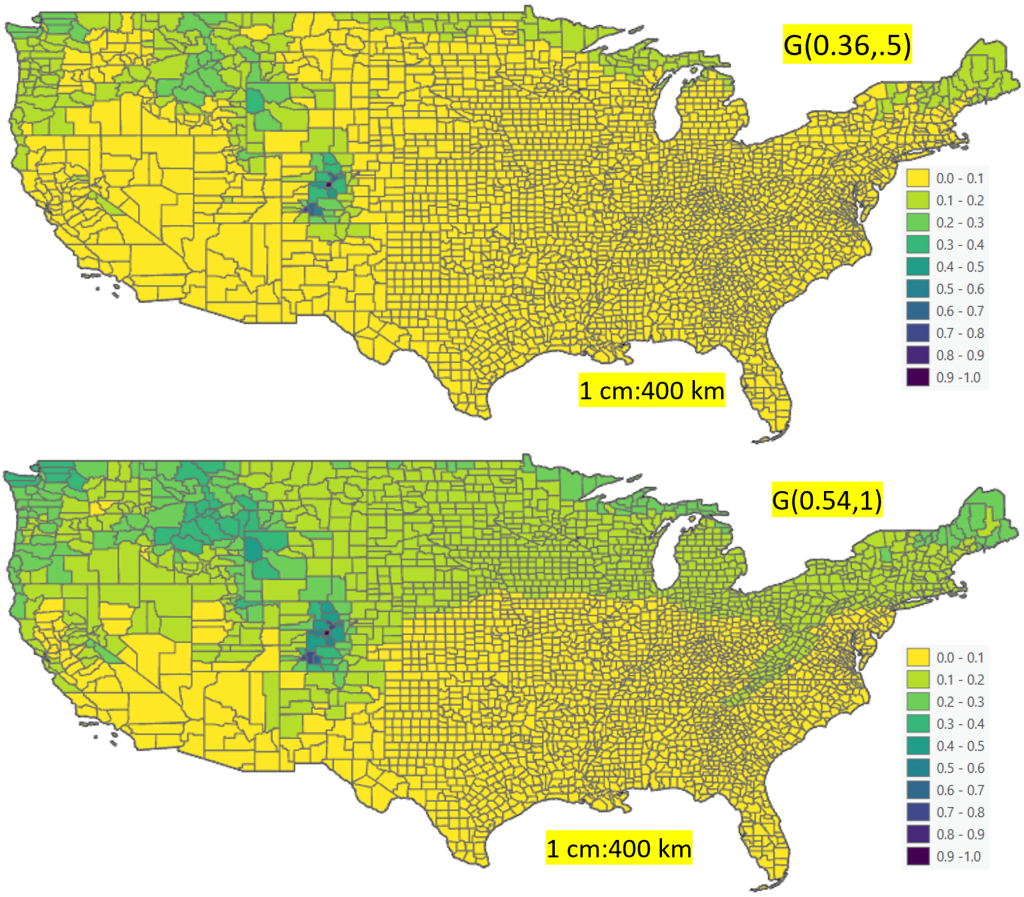


Figure B.17: The overall suitability for each county in the USA for the invasive EAB colonisation under the model 17 ($G^D_{\alpha_1} K_{\exp}$) and model 16 ($N^D_{\alpha_1,.86} K_{\exp}$), respectively. Each of these model is the spatio-temporal colonisation–dispersal model based on estimated ash density and the MAGDD. The value calculated for each county is the mean of the posterior distribution of the suitability function.G21 G24

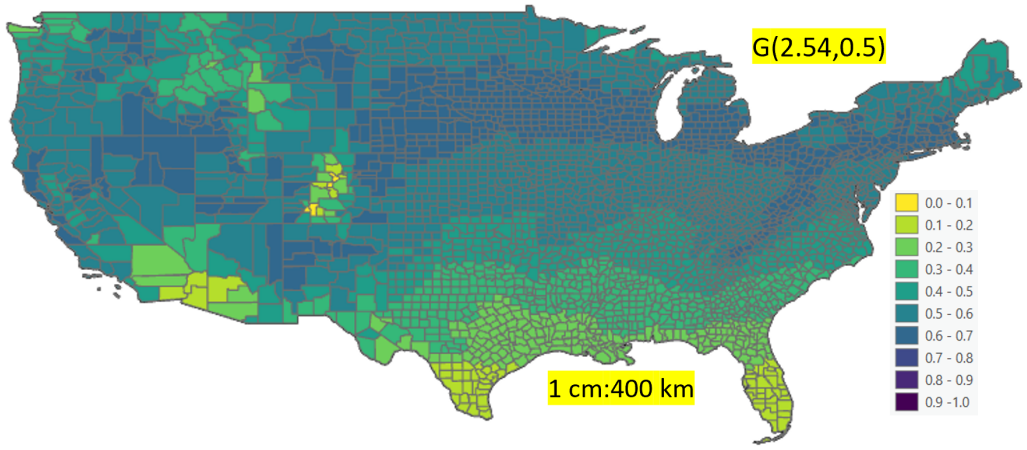


Figure B.18: The overall suitability for each county in the USA for the invasive EAB colonisation under the model 15 ($G^{E}_{\alpha_1,0.5}K_{\exp}$). Each of these models is the spatio-temporal colonisation–dispersal model based on estimated ash density and the MAGDD. The value calculated for each county is the mean of the posterior distribution of the suitability function.

Bibliography

[1] Food and Agriculture Organization of the United Nations (FAO), "Global Action for Fall Armyworm Control," accessed on March 24, 2023. [Online]. Available: https://www.fao.org/fall-armyworm/monitoring-tools/faw-map/en/

[2] H. C. Baladhiya, D. B. Sisodiya, and N. P. Pathan, "A review on pink stem borer, Sesamia inferens Walker: A threat to cereals," *Journal of Entomology and Zoology Studies*, vol. 6, no. 3, pp. 1235–1239, 2018.

[3] P. Kumar, J. Kaur, S. B. Suby, J. C. Sekhar, and S. P. Lakshmi, *Pests and Their Management*. Singapore: Springer Singapore, 2018, ch. Pests of Maize, pp. 51–79. [Online]. Available: https://doi.org/10.1007/978-981-10-8687-8$_3$

[4] K. Viswajyothi, N. Aggarwal, and J. Jindal, "The biology of Sesamia inferens (Walker)(*Lepidoptera: Noctuidae*) on maize in the north western plains of India," *Acta Phytopathologica et Entomologica Hungarica*, vol. 54, no. 1, pp. 69–84, 2019.

[5] J. N. Mandrekar, "Receiver operating characteristic curve in diagnostic test assessment," *Journal of Thoracic Oncology*, vol. 5, no. 9, pp. 1315–1316, 2010.

[6] S. Niassy, S. Ekesi, L. Migiro, and W. Otieno, *Introduction: An Overview of the Impacts of Invasive Insect Species on Agriculture*. Cham: Springer International Publishing, 2020, pp. 1–9. [Online]. Available: https://doi.org/10.1007/978-3-030-41083-4$_1$

[7] A. Ricciardi, *Invasive Species*. New York, NY: Springer New York, 2012, pp. 5547–5560. [Online]. Available: https://doi.org/10.1007/978-1-4419-0851-3$_5$74

[8] C. J. Crous, T. I. Burgess, J. J. Le Roux, D. M. Richardson, B. Slippers, and M. J. Wingfield, "Ecological disequilibrium drives insect pest and pathogen accumulation in non-native trees," *AoB Plants*, vol. 9, no. 1, p. plw081, 2017.

[9] IUCN, "IUCN Standard to support global action on invasive alien species," Sep. 2020, accessed: 2023-07-14. [Online]. Available: https://www.iucn.org/news/species/202009/iucn-standard-support-global-action-invasive-alien-species

[10] R. Early, B. A. Bradley, J. S. Dukes, J. J. Lawler, J. D. Olden, D. M. Blumenthal, P. Gonzalez, E. D. Grosholz, I. Ibañez, L. P. Miller *et al.*, "Global threats from invasive alien species in the twenty-first century and national response capacities," *Nature Communications*, vol. 7, p. 12485, 2016.

[11] M. Tokuda and N. Uechi, *Invasive Species*. Singapore: Springer Singapore, 2021, ch. Invasive Species, pp. 255–267. [Online]. Available: https://doi.org/10.1007/978-981-33-6534-6$_1$2

[12] A. A. Batabyal and D. M. Lee, "The infinitesimal, the deterministic, and the probabilistic: alternate container inspection policies in invasive species management," *Biological Invasions*, vol. 8, pp. 1663–1671, 2006.

[13] A. M. Liebhold, E. G. Brockerhoff, L. J. Garrett, J. L. Parke, and K. O. Britton, "Live Plant Imports: the Major Pathway For Forest Insect And Pathogen Invasions Of The Us," *Frontiers in Ecology and the Environment*, vol. 10, pp. 135–143, 2012.

[14] V. Valenta, D. Moser, S. Kapeller, and F. Essl, "A new forest pest in Europe: a review of Emerald ash borer (*Agrilus planipennis*) invasion," *Journal of Applied Entomology*, vol. 141, no. 7, pp. 507–526, 2016.

[15] A. E. Mayfield, S. J. Seybold, W. R. Haag, M. T. Johnson, B. K. Kerns, J. C. Kilgo, D. J. Larkin, R. D. Lucardi, B. D. Moltzan, D. E. Pearson, J. D. Rothlisberger, J. D. Schardt, M. K. Schwartz, and M. K. Young, *Impacts of Invasive Species in Terrestrial and Aquatic Systems in the United States*. Cham: Springer International Publishing, 2021, pp. 5–39. [Online]. Available: https://doi.org/10.1007/978-3-030-45367-1$_2$

[16] M. Kenis, M.-A. Auger-Rozenberg, A. Roques, L. Timms, C. Péré, M. J. W. Cock, J. Settele, S. Augustin, and C. Lopez-Vaamonde, "Ecological effects of invasive alien insects," *Biological Invasions*, vol. 11, pp. 21–45, 2009.

[17] N.-C. Freistetter, G. S. Simmons, Y. Wu, D. C. Finger, and R. Hood-Nowotny, "Tracking global invasion pathways of the spongy moth (*Lepidoptera: Erebidae*) to the

United States using stable isotopes as endogenous biomarkers," *Ecology and Evolution*, vol. 12, no. 7, p. e9092, 2022.

[18] P. Brown and H. Roy, "Native Ladybird Decline Caused By the Invasive Harlequin Ladybird harmonia Axyridis : Evidence From A Long-term Field Study," *Insect Conserv Divers*, vol. 11, pp. 230–239, 2017.

[19] M. P. Hill, B. Gallardo, and J. S. Terblanche, "A Global Assessment of Climatic Niche Shifts and Human Influence in Insect Invasions," *Global Ecology and Biogeography*, 2017.

[20] C. J. A. Bradshaw, B. Leroy, C. Bellard, D. Roiz, C. Albert, A. Fournier, M. Barbet-Massin, J.-P. Salles, F. Simard, and F. Courchamp, "Massive yet grossly underestimated global costs of invasive insects," *Nature communications*, vol. 7, no. 1, p. 12986, 2016.

[21] H. De Groote, S. C. Kimenju, B. Munyua, S. Palmas, M. Kassie, and A. Bruce, "Spread and impact of fall armyworm (*Spodoptera frugiperda* J.E. Smith) in maize production areas of Kenya," *Agriculture, Ecosystems Environment*, vol. 292, p. 106804, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0167880919304219

[22] A. M. Simmons and B. Wiseman, "James edward smith-taxonomic author of the fall armyworm," *Florida Entomologist*, pp. 271–276, 1993.

[23] N. Gupta, S. Verma, P. Sharma, M. Thakur, P. Sharma, and D. Devi, "Status of invasive insect pests of India and their natural enemies," 2019.

[24] S. Rakshit, C. R. Ballal, Y. G. Prasad, J. C. Sekhar, P. Lakshmi Soujanya, S. B. Suby, S. L. Jat, G. Siva Kumar, and J. V. Prasad, "Fight against fall armyworm *Spodoptera frugiperda* (JE Smith)," 2019.

[25] R. A. Gutierrez-Moreno, *Susceptibility of the fall armyworm,* Spodoptera frugiperda *(JE Smith)(Lepidoptera: Noctuidae) to Bacillus thuringiensis (Bt) proteins and synthetic insecticides from different corn production systems in Mexico and Puerto Rico*. Michigan State University, 2017.

[26] S. Deshmukh, H. B. Pavithra, C. M. Kalleshwaraswamy, B. K. Shivanna, M. S. Maruthi, and D. Mota-Sanchez, "Field Efficacy of Insecticides for Management of Invasive Fall

Armyworm, *Spodoptera frugiperda* (J. E. Smith) (Lepidoptera: Noctuidae) on Maize in India," *The Florida entomologist*, vol. 103, no. 2, pp. 221–227, 2020.

[27] L. Bengyella, B. A. Hetsa, D. J. Fonmboh, and R. C. Jose, "Assessment of damage caused by evolved fall armyworm on native and transgenic maize in South Africa," *Phytoparasitica*, vol. 49, no. 1, pp. 1–12, 2021.

[28] P. Luginbill, *The fall army worm.* US Department of Agriculture, 1928, no. 34.

[29] S. B. Suby, P. L. Soujanya, P. Yadava, J. Patil, K. Subaharan, G. S. Prasad, K. S. Babu, S. L. Jat, K. R. Yathish, J. Vadassery, V. K. Kalia, N. Bakthavatsalam, J. C. Shekhar, and S. Rakshit, "Invasion of Fall Armyworm (*Spodoptera frugiperda*) in India: Nature, Distribution, Management and Potential Impact," *Current Science*, vol. 119, no. 1, pp. 44–51, 2020.

[30] A. Sangomla and I. Kukreti, "Fall Armyworm attack: The damage done," https://www.downtoearth.org.in/coverage/agriculture/fall-armyworm-attack-the-damage-done-63445, 2019, accessed: 2020-05-06.

[31] J. L. Maino, R. Schouten, K. Overton, R. Day, S. Ekesi, B. Bett, M. Barton, P. C. Gregg, P. A. Umina, and O. L. Reynolds, "Regional and seasonal activity predictions for fall armyworm in australia," *Current Research in Insect Science*, vol. 1, p. 100010, 2021.

[32] N. Nayyar, R. G. Gracy, T. R. Ashika, G. Mohan, R. S. Swathi, M. Mohan, M. Chaudhary, N. Bakthavatsalam, and T. Venkatesan, "Population structure and genetic diversity of invasive Fall Armyworm after 2 years of introduction in India," *Scientific reports*, vol. 11, no. 1, p. 7760, 2021.

[33] "Fall armyworm, *Spodoptera frugiperda* (JE Smith)(Insecta: Lepidoptera: Noctuidae), author=Capinera, John L, journal=EDIS, volume=2002, number=7, year=2002."

[34] Z. Abro, E. Kimathi, H. De Groote, T. Tefera, S. Sevgan, S. Niassy, and M. Kassie, "Socioeconomic and health impacts of fall armyworm in Ethiopia," *PLoS One*, vol. 16, no. 11, p. e0257736, 2021.

[35] S. Deshmukh, H. B. Pavithra, C. M. Kalleshwaraswamy, B. K. Shivanna, M. S. Maruthi, and D. Mota-Sanchez, "Field Efficacy of Insecticides for Management of Invasive Fall

Armyworm, *Spodoptera frugiperda* (J. E. Smith) (Lepidoptera: Noctuidae) on Maize in India," *The Florida entomologist*, vol. 103, no. 2, pp. 221–227, 2020.

[36] C. Parihar, S. Jat, A. Singh, R. S. Kumar, K. Hooda, C. GK, and D. Singh, "From Directorâs Desk," https://iimr.icar.gov.in/from-directors-desk/, 2021, directorate of Maize Research (Indian Council of Agricultural Research).

[37] A. Madhukar, V. Kumar, and K. Dashora, "Spatial and Temporal Trends in the Yields of Three Major Crops: Wheat, Rice and Maize in India," *International Journal of Plant Production*, pp. 1–21, 2019.

[38] N. Singh, P. Sharma, and M. Kamboj, "Maize scenario in haryana: A brief review," *Int. J. Pure App. Biosci*, vol. 5, no. 6, pp. 1616–1623, 2017.

[39] C. Parihar, S. Jat, A. Singh, R. S. Kumar, K. Hooda, C. GK, and D. Singh, "Maize production technologies in India," www.maizeindia.org, 2011, directorate of Maize Research (Indian Council of Agricultural Research).

[40] R. Kumar, K. Srinivas, and N. Sivaramane, "Assessment of the maize situation, outlook and investment opportunities in India," *Country Report–Regional Assessment Asia (MAIZE-CRP), National Academy of Agricultural Research Management, Hyderabad, India*, vol. 133, 2013.

[41] V. A. Keenan, G. Marion, and A. Kleczkowski, "Modelling framework for invasive pests: emerald ash borer as a case study. Project Final Report." Scotland's Centre of Expertise for Plant Health (PHC)., Tech. Rep., 2020.

[42] D. Cappaert, D. G. McCullough, T. M. Poland, and N. W. Siegert, "Emerald ash borer in north america: a research and regulatory challenge," *American Entomologist. 51 (3): 152-165.*, vol. 51, no. 3, 2005.

[43] J. R. Muirhead, B. Leung, C. van Overdijk, D. W. Kelly, K. Nandakumar, K. R. Marchant, and H. J. MacIsaac, "Modelling local and long-distance dispersal of invasive emerald ash borer *Agrilus planipennis* (Coleoptera) in North America," *Diversity and Distributions*, vol. 12, no. 1, pp. 71–79, 2006.

[44] N. A. Straw, D. T. Williams, O. Kulinich, and Y. I. Gninenko, "Distribution, impact and rate of spread of emerald ash borer *Agrilus planipennis* (Coleoptera: Buprestidae) in the Moscow region of Russia," *Forestry*, vol. 86, no. 5, pp. 515–522, 2013.

[45] L. Liang and S. Fei, "Divergence of the potential invasion range of emerald ash borer and its host distribution in North America under climate change," *Climatic Change*, vol. 122, no. 4, pp. 735–746, 2014.

[46] M. J. Orlova-Bienkowskaja and A. O. Bieńkowski, "Modeling long-distance dispersal of emerald ash borer in european russia and prognosis of spread of this pest to neighboring countries within next 5 years," *Ecology and evolution*, vol. 8, no. 18, pp. 9295–9304, 2018.

[47] T. M. Poland *et al.*, "Twenty million ash trees later: current status of emerald ash borer in Michigan," *Newsletter of the Michigan entomological society*, vol. 52, no. 1&2, pp. 10–14, 2007.

[48] A. Juliann E., L. Brian, K. Kent, C. Corey, B. Kerry O, E. Jeffrey, F. Susan J., H. Robert G., H. Thomas P., L. Andrew M., M. Deborah G., and H. Betsy Von, "Economic impacts of non-native forest insects in the continental United States," *PLoS one*, vol. 6, no. 9, p. e24587, 2011.

[49] D. A. Herms, D. G. McCullough *et al.*, "Emerald ash borer invasion of North America: history, biology, ecology, impacts, and management," *Annual review of entomology*, vol. 59, no. 1, pp. 13–30, 2014.

[50] M. J. Orlova-Bienkowskaja and A. O. Bieńkowski, "The life cycle of the emerald ash borer *Agrilus planipennis* in European Russia and comparisons with its life cycles in Asia and North America," *Agricultural and Forest Entomology*, vol. 18, no. 2, pp. 182–188, 2016.

[51] T. M. Poland, Y. Chen, J. Koch, and D. Pureswaran, "Review of the emerald ash borer (coleoptera: Buprestidae), life history, mating behaviours, host plant selection, and host resistance," *The Canadian Entomologist*, vol. 147, no. 3, pp. 252–262, 2015.

[52] L. M. Moley, D. G. Goodin, and W. P. Winslow, "Leaf-Level Spectroscopy for Analysis of Invasive Pest Impact on Trees in a Stressed Environment: An Example Using Emerald Ash Borer (*Agrilus planipennis* Fairmaire) in Ash Trees (Fraxinus spp.), Kansas, USA," *Environments*, vol. 9, no. 4, 2022. [Online]. Available: https://www.mdpi.com/2076-3298/9/4/42

[53] J. J. Duan, K. J. Abell, L. S. Bauer, J. Gould, and R. Van Driesche, "Natural enemies implicated in the regulation of an invasive pest: a life table analysis of the population

dynamics of the emerald ash borer," *Agricultural and Forest Entomology*, vol. 16, pp. 406–416, 2014.

[54] L. S. Bauer, R. A. Haack, D. L. Miller, T. R. Petrice, and H. Liu, "Emerald ash borer life cycle," *Mastro and Reardon*, p. 8, 2004.

[55] N. W. Siegert, R. J. Mercader, and D. G. McCullough, "Spread and dispersal of emerald ash borer (coleoptera: Buprestidae): estimating the spatial dynamics of a difficult-to-detect invasive forest pest," *The Canadian Entomologist*, vol. 147, no. 3, pp. 338–348, 2015.

[56] K. Cuddington, S. Sobek-Swant, J. C. Crosthwaite, D. B. Lyons, and B. J. Sinclair, "Probability of emerald ash borer impact for canadian cities and north america: a mechanistic model," *Biological Invasions*, vol. 20, no. 9, pp. 2661–2677, 2018.

[57] Emerald Ash Borer Information, "Emerald Ash Borer Information Network," http://www.emeraldashborer.info/, 2023, accessed: 04 January 2023.

[58] K. F. Kovacs, R. G. Haight, D. G. McCullough, R. J. Mercader, N. W. Siegert, and A. M. Liebhold, "Cost of potential emerald ash borer damage in u.s. communities, 2009â2019," *Ecological Economics*, vol. 69, no. 3, pp. 569–578, 2010. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0921800909003681

[59] T. D. Sydnor, M. Bumgardner, A. Todd *et al.*, "The potential economic impacts of emerald ash borer (*Agrilus planipennis*) on Ohio, US, communities," *Arboriculture and Urban Forestry*, vol. 33, no. 1, p. 48, 2007.

[60] D. G. McCullough, "Challenges, tactics and integrated management of emerald ash borer in North America," *Forestry: An International Journal of Forest Research*, vol. 93, no. 2, pp. 197–211, 2020.

[61] L. S. Bauer, H. Liu, R. A. Haack, R. Gao, T. Zhao, D. L. Miller, T. R. Petrice, V. Mastro, and R. Reardon, "Update on emerald ash borer natural enemy surveys in Michigan and China," *Mastro and Reardon*, pp. 71–72, 2005.

[62] J. J. Duan, J. R. Gould, N. F. Quinn, T. R. Petrice, B. H. Slager, T. M. Poland, L. S. Bauer, C. E. Rutledge, J. S. Elkinton, and R. G. Van Driesche, "Protection of North American ash against emerald ash borer with biological control: ecological premises and progress toward success," *BioControl*, vol. 68, no. 2, pp. 87–100, 2023.

[63] D. Cappaert, D. G. McCullough *et al.*, "Occurrence and seasonal abundance of *Atanycolus cappaerti* (Hymenoptera: Braconidae) a native parasitoid of emerald ash borer, *Agrilus plonipennis* (Coleoptera: Buprestidae)." *Great Lakes Entomologist*, vol. 42, no. 1/2, pp. 16–29, 2009.

[64] L. E. Roscoe, D. B. Lyons, and S. M. Smith, "Observations on the life-history traits of the North American parasitoid *Phasgonophora sulcata* Westwood (Hymenoptera: Chalcididae) attacking *Agrilus planipennis* (Coleoptera: Buprestidae) in Ontario, Canada," *The Canadian Entomologist*, vol. 148, no. 3, pp. 294–306, 2016.

[65] D. G. McCullough, T. M. Poland, A. C. Anulewicz, P. Lewis, and D. Cappaert, "Evaluation of *Agrilus planipennis* (Coleoptera: Buprestidae) control provided by emamectin benzoate and two neonicotinoid insecticides, one and two seasons after treatment," *Journal of economic entomology*, vol. 104, no. 5, pp. 1599–1612, 2011.

[66] R. J. Mercader, D. G. McCullough, A. J. Storer, J. M. Bedford, R. Heyd, N. W. Siegert, S. Katovich, and T. M. Poland, "Estimating local spread of recently established emerald ash borer, *Agrilus planipennis*, infestations and the potential to influence it with a systemic insecticide and girdled ash trees," *Forest Ecology and Management*, vol. 366, pp. 87–97, 2016. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0378112716300160

[67] M. Ghahramani, R. Karimzadeh, S. Iranipour, and A. Sciarretta, "Does harvesting affect the spatio-temporal signature of pests and natural enemies in alfalfa fields?" *Agronomy*, vol. 9, no. 9, p. 532, 2019.

[68] A. Siddiqua, M. A. Kabir, T. Ferdous, I. B. Ali, and L. A. Weston, "Evaluating plant disease detection mobile applications: Quality and limitations," *Agronomy*, vol. 12, no. 8, p. 1869, 2022.

[69] M. Pruvot, W. Hutchins, and K. Orsel, "Statistical evaluation of a commercial Neospora caninum competitive ELISA in the absence of a gold standard: application to wild elk (*Cervus elaphus*) in Alberta," *Parasitology research*, vol. 113, pp. 2899–2905, 2014.

[70] V. Marmara, A. Cook, and A. Kleczkowski, "Estimation of force of infection based on different epidemiological proxies: 2009/2010 influenza epidemic in malta," *Epidemics*, vol. 9, pp. 52–61, 2014.

[71] N. Petrellis, "Plant disease diagnosis for smart phone applications with extensible set of diseases," *Applied Sciences*, vol. 9, no. 9, p. 1952, 2019.

[72] H. E. Roy, M. J. Pocock, C. D. Preston, D. B. Roy, J. Savage, J. Tweddle, and L. Robinson, "Understanding citizen science and environmental monitoring: final report on behalf of uk environmental observation framework," 2012.

[73] R. C. Jordan, H. L. Ballard, and T. B. Phillips, "Key issues and new approaches for evaluating citizen-science learning outcomes," *Frontiers in Ecology and the Environment*, vol. 10, no. 6, pp. 307–309, 2012.

[74] G. Silva, J. Tomlinson, N. Onkokesung, S. Sommer, L. Mrisho, J. Legg, I. P. Adams, Y. Gutierrez-Vazquez, T. P. Howard, A. Laverick, O. Hossain, Q. Wei, K. M. Gold, and N. Boonham, "Plant pest surveillance: From satellites to molecules," *Emerging topics in life sciences*, vol. 5, no. 2, pp. 275–287, 2021.

[75] E. Crocker, B. Condon, A. Almsaeed, B. Jarret, C. D. Nelson, A. G. Abbott, D. Main, and M. Staton, "TreeSnap: A citizen science app connecting tree enthusiasts and forest scientists," *PLANTS, PEOPLE, PLANET*, vol. 2, no. 1, pp. 47–52, jan 2020. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/ppp3.41

[76] J. Z. Barbosa, S. A. Prior, G. Q. Pedreira, A. C. V. Motta, G. C. Poggere, and G. D. Goularte, "Global trends in apps for agriculture," *Multi-Science Journal*, vol. 3, no. 1, p. 16, apr 2020.

[77] D. G. Panpatte, "Artificial intelligence in agriculture: An emerging era of research," *Anand Agricultural University*, 2018.

[78] S. E. E. Inwood and V. H. Dale, "State of apps targeting management for sustainability of agricultural landscapes. a review," *Agronomy for sustainable development*, vol. 39, no. 1, p. 8, 2019.

[79] S. Wang, S. Di Tommaso, J. Faulkner, T. Friedel, A. Kennepohl, R. Strey, and D. B. Lobell, "Mapping crop types in southeast india with smartphone crowdsourcing and deep learning," *Remote Sensing*, vol. 12, no. 18, p. 2957, 2020.

[80] A. C. Hampf, C. Nendel, S. Strey, and R. Strey, "Biotic Yield Losses in the Southern Amazon, Brazil: Making Use of Smartphone-Assisted Plant Disease Diagnosis Data," *Frontiers in Plant Science*, vol. 12, p. 548, 2021.

[81] "Plantix on networked," https://plantix.net/en/blog/plantix-networked-india, 2016, accessed: 2021-05-17.

[82] Y. Y. Lasania, "Plantix on livemint," https://plantix.net/en/blog/plantix-livemint-india, 2017, accessed: 2021-05-18.

[83] "From Berlin to Indore: Why did one of the top German female cofounders and CEOs relocate to India?" https://plantix.net/en/blog/simone-strey-moves-from-berlin-to-india, 2021, accessed: 2021-05-16.

[84] "Plantix on India Agro Net," https://plantix.net/en/blog/plantix-india-agro-net-india, 2016, accessed: 2021-05-17.

[85] "Plantix on the hindu," https://plantix.net/en/blog/plantix-hindu-india-telangana-2016, 2016, accessed: 2021-05-17.

[86] "MOBILE APP TO HELP FARMERS OVERCOME CROP DAMAGE LAUNCHED IN INDIA," http://www.icrisat.org/mobile-app-to-help-farmers-overcome-crop-damage-launched-in-india/, 2017, accessed: 2021-05-17.

[87] D. J. Kotze, R. B. OâHara, and S. Lehvävirta, "Dealing with varying detection probability, unequal sample sizes and clumped distributions in count data," *Plos One*, vol. 7, no. 7, p. e40923, 2012.

[88] T. J. Bird, A. E. Bates, J. S. Lefcheck, N. A. Hill, R. J. Thomson, G. J. Edgar, R. D. Stuart-Smith, S. Wotherspoon, M. Krkosek, J. F. Stuart-Smith *et al.*, "Statistical solutions for error and bias in global citizen science datasets," *Biological Conservation*, vol. 173, pp. 144–154, 2014.

[89] T. Bardhan, T. Saklani, P. Jalal, J. Barakoti, V. Kameswari *et al.*, "Science behind user friendliness of agricultural mobile apps: A study on readability," *Indian Journal of Agricultural Sciences*, vol. 92, no. 1, pp. 55–8, 2022.

[90] M. Kumar and L. Agrawal, "Empowering farming community through mobile applications: Changing scenarios," *International Journal of Scientific & Technology Research*, vol. 9, no. 3, pp. 58–61, 2020.

[91] O. Akinyemi *et al.*, "Exploring A Mobile Application for Pest and Disease Symptomatic Diagnosis," *Journal of Agricultural Science*, vol. 12, no. 5, pp. 1–12, 2023, study on the accuracy of the Plantix app in diagnosing fall armyworm damage in maize.

[92] C. M. Umemneku Chikere, K. Wilson, S. Graziadio, L. Vale, and A. J. Allen, "Diagnostic test evaluation methodology: A systematic review of methods employed to evaluate diagnostic tests in the absence of gold standard â An update," *PLOS ONE*, vol. 14, no. 10, p. e0223832, oct 2019. [Online]. Available: https://doi.org/10.1371/journal.pone.0223832

[93] M. Martínez-García, S. García-Gutierrez, L. Barreñada Taleb, R. Armañanzas, I. Inza, and J. A. Lozano, "Clinical severity prediction of COVID-19 admitted patients in Spain: SEMI and REDISSEC cohorts," *medRxiv*, 2023.

[94] M. Opsteegh, P. Teunis, M. Mensink, L. Züchner, A. Titilincu, M. Langelaar, and J. van der Giessen, "Evaluation of elisa test characteristics and estimation of toxoplasma gondii seroprevalence in dutch sheep using mixture models," *Preventive veterinary medicine*, vol. 96, no. 3-4, pp. 232–240, 2010.

[95] A. M. Prasad, L. R. Iverson, M. P. Peters, J. M. Bossenbroek, S. N. Matthews, T. D. Sydnor, and M. W. Schwartz, "Modeling the risk of invasion and spread of the Emerald Ash Borer (*Agrilus planipennis*) in Ohio and prediction of its impact on Ohio's forests," *Landscape ecology*, vol. 25, no. 3, pp. 353–369, 2009.

[96] S. F. Ward, A. M. Liebhold, R. S. Morin, and S. Fei, "Temporal dynamics and drivers of landscape-level spread by emerald ash borer," *Journal of Applied Ecology*, vol. 57, no. 6, pp. 1165–1176, 2020.

[97] M. J. Orlova-Bienkowskaja and A. O. BieÅkowski, "Low heat availability could limit the potential spread of the emerald ash borer to northern europe (prognosis based on growing degree days per year)," *Insects (Basel, Switzerland)*, vol. 13, no. 1, p. 52, 2022.

[98] A. M. Prasad, L. R. Iverson, M. P. Peters, and et al., "Modeling the Invasive Emerald Ash Borer Risk of Spread using a Spatially Explicit Cellular Model," *Landscape Ecology*, vol. 25, pp. 353–369, 2010. [Online]. Available: https://doi.org/10.1007/s10980-009-9434-9

[99] C. R. Webb, T. Mona, and C. A. Gilligan, "Predicting the potential for spread of emerald ash borer (*Agrilus planipennis*) in Great Britain: What can we learn from other affected areas?" *Plants, People, Planet*, vol. 3, no. 4, pp. 402–413, 2021.

[100] R. J. Mercader, A. M. Siegert, A. M. Liebhold, and D. G. McCullough, "Dispersal of the emerald ash borer, *Agrilus planipennis*, in newly-colonized sites," *Agricultural and Forest Entomology. 11: 421? 424.*, vol. 11, 2009.

[101] A. M. Prasad, L. R. Iverson, M. P. Peters, J. M. Bossenbroek, S. N. Matthews, T. Davis Sydnor, and M. W. Schwartz, "Modeling the invasive emerald ash borer risk of spread using a spatially explicit cellular model," *Landscape ecology*, vol. 25, no. 3, pp. 353–369, 2010.

[102] S. Sobek-Swant, D. A. Kluza, K. Cuddington, and D. B. Lyons, "Potential distribution of emerald ash borer: What can we learn from ecological niche models using Maxent and GARP?" *Forest Ecology and Management*, vol. 281, pp. 23–31, 2012.

[103] S. Catterall, A. R. Cook, G. Marion, A. Butler, and P. E. Hulme, "Accounting for uncertainty in colonisation times: a novel approach to modelling the spatio-temporal dynamics of alien invasions using distribution data," *Ecography (Copenhagen)*, vol. 35, no. 10, pp. 901–911, 2012.

[104] F. Fischer, D. Zocholl, G. Rauch, B. Levis, A. Benedetti, B. Thombs, M. Rose, and P. Kostoulas, "Prevalence estimates of major depressive disorder in 27 European countries from the European Health Interview Survey: accounting for imperfect diagnostic accuracy of the PHQ-8," *BMJ Ment Health*, vol. 26, no. 1, 2023.

[105] R. Parikh, A. Mathai, S. Parikh, G. C. Sekhar, and R. Thomas, "Understanding and using sensitivity, specificity and predictive values," *Indian journal of ophthalmology*, vol. 56, no. 1, p. 45, 2008.

[106] A. Albert, "Biostatistics: Facing the Interpretation of 2×2 Tables," *Journal of the Belgian Society of Radiology*, vol. 101, no. Suppl 2, 2017.

[107] B. Ghojogh, A. Ghojogh, M. Crowley, and F. Karray, "Fitting a mixture distribution to data: tutorial," *arXiv preprint arXiv:1901.06708*, 2019.

[108] T. Hastie, R. Tibshirani, and J. Friedman, *The elements of statistical learning: data mining, inference, and prediction.* Springer Science & Business Media, 2009.

[109] T. K. Moon, "The expectation-maximization algorithm," *IEEE Signal processing magazine*, vol. 13, no. 6, pp. 47–60, 1996.

[110] I. J. Myung, "Tutorial on maximum likelihood estimation," *Journal of mathematical Psychology*, vol. 47, no. 1, pp. 90–100, 2003.

[111] C. B. Do and S. Batzoglou, "What is the expectation maximization algorithm?" *Nature biotechnology*, vol. 26, no. 8, pp. 897–899, 2008.

[112] K. Hajian-Tilaki, "Receiver operating characteristic (roc) curve analysis for medical diagnostic test evaluation," *Caspian journal of internal medicine*, vol. 4, no. 2, p. 627, 2013.

[113] R. Fluss, D. Faraggi, and B. Reiser, "Estimation of the Youden Index and its associated cutoff point," *Biometrical Journal: Journal of Mathematical Methods in Biosciences*, vol. 47, no. 4, pp. 458–472, 2005.

[114] J. A. Hanley and B. J. McNeil, "The meaning and use of the area under a receiver operating characteristic (ROC) curve." *Radiology*, vol. 143, no. 1, pp. 29–36, 1982.

[115] A. M. Carrington, P. W. Fieguth, H. Qazi, and et al., "A new concordant partial AUC and partial c statistic for imbalanced data in the evaluation of machine learning algorithms," *BMC Medical Informatics and Decision Making*, vol. 20, no. 4, pp. 1–10, 2020. [Online]. Available: https://doi.org/10.1186/s12911-019-1014-6

[116] G. S. McMaster and W. W. Wilhelm, "Growing degree-days: one equation, two interpretations," *Agricultural and Forest Meteorology*, vol. 87, no. 4, pp. 291–300, 1997. [Online]. Available: https://doi.org/10.1016/S0168-1923(97)00027-0

[117] P. Choidis, A. Sharma, G. Grottesi, and D. Kraniotis, "Climate change impact on the degradation of historically significant wooden furniture in a cultural heritage building in Vestfold, Norway," *E3S Web of Conferences*, 2022.

[118] G. Zhou and Q. Wang, "A Nonlinear Method for Calculating Growing Degree Days," *Scientific Reports*, vol. 8, no. 1, pp. 1–9, 2018, this article is licensed under a Creative Commons Attribution 4.0 International License. [Online]. Available: https://www.nature.com/articles/s41598-018-30110-1

[119] T. Simisky, "Growing Degree Days for Management of Insect Pests in the Landscape," 2017, accessed:28 August 2023. [Online]. Available: https://ag.umass.edu/landscape/fact-sheets/growing-degree-days-for-management-of-insect-pests-in-landscape

[120] N. E. Adams, "Using Growing Degree Days For Insect Management," *University of New Hampshire, Cooperative Extension*, vol. 1, no. 4, 2013.

[121] A. Cook, G. Marion, A. Butler, and G. Gibson, "Bayesian inference for the spatio-temporal invasion of alien species," *Bulletin of mathematical biology*, vol. 69, no. 6, pp. 2005–2025, 2007.

[122] D. Van Ravenzwaaij, P. Cassey, and S. D. Brown, "A simple introduction to markov chain monte–carlo sampling," *Psychonomic bulletin & review*, vol. 25, no. 1, pp. 143–154, 2018.

[123] G. Marian, D. Walker, D. Swain, M. Hutchings, and A. Cook, *Towards an integrated approach to stochastic process-based modelling: with applications to animal behaviour and spatio-temporal spread.* United Kingdom: CAB International, 2007, p. 231.

[124] P. J. Green, K. Łatuszyński, M. Pereyra, and C. P. Robert, "Bayesian computation: a perspective on the current state, and sampling backwards and forwards," *arXiv preprint arXiv:1502.01148*, 2015.

[125] C. Andrieu and J. Thoms, "A tutorial on adaptive MCMC," *Statistics and computing*, vol. 18, pp. 343–373, 2008.

[126] V. Roy, "Convergence diagnostics for markov chain monte carlo," *Annual Review of Statistics and Its Application*, vol. 7, pp. 387–412, 2020.

[127] N. Toft, G. T. Innocent, G. Gettinby, and S. W. J. Reid, "Assessing the convergence of Markov Chain Monte Carlo methods: An example from evaluation of diagnostic tests in absence of a gold standard," *Preventive Veterinary Medicine*, vol. 79, no. 2, pp. 244–256, 2007. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0167587707000037

[128] H. Du, Z. Ke, G. Jiang, and S. Huang, "The Performances of Gelman-Rubin and Geweke's Convergence Diagnostics of Monte Carlo Markov Chains in Bayesian Analysis," *Journal of Behavioral Data Science*, vol. 2, no. 2, pp. 47–72, 2022.

[129] M. Plummer, N. Best, K. Cowles, and K. Vines, "Package âcodaâ," *URL http://cran. r-project. org/web/packages/coda/coda. pdf, accessed January*, vol. 25, p. 2015, 2015. [Online]. Available: https://cran.r-project.org/web/packages/coda/

[130] J. Guo, J. Gabry, B. Goodrich, and S. Weber, *Package âărstanâ*, 2020, accessed: 2023-01-15. [Online]. Available: https://pj.freefaculty.org/guides/crmda$_w$*orkshops/sem/Archive/sem* − 4/*literature/manuals/rstan.pdf*

[131] Y. Huang, H. Lv, Y. Dong, W. Huang, G. Hu, Y. Liu, H. Chen, Y. Geng, J. Bai, P. Guo *et al.*, "Mapping the Spatio-Temporal Distribution of Fall Armyworm in China by Coupling Multi-Factors," *Remote Sensing*, vol. 14, no. 17, p. 4415, 2022.

[132] R. Epanchin-Niell, A. L. Thompson, and T. Treakle, "Public contributions to early detection of new invasive pests," *Conservation Science and Practice*, vol. 3, no. 6, p. e422, 2021.

[133] F. . R. A. Department for Environment, "UK Plant Health Risk Register," https://planthealthportal.defra.gov.uk/pests-and-diseases/uk-plant-health-risk-register/, 2014, 05/09/2022.

[134] D. P. S. Badwal, R. Kumar, V. Singh, S. Kaur, and M. Kumar, "Threat of fall armyworm in India and it's management- A Review," *Journal of Entomology and Zoology Studies*, vol. 8, no. 4, pp. 19–22, 2020.

[135] O. Navik, A. N. Shylesha, J. Patil, T. Venkatesan, Y. Lalitha, and T. R. Ashika, "Damage, distribution and natural enemies of invasive fall armyworm *Spodoptera frugiperda* (JE smith) under rainfed maize in Karnataka, India," *Crop Protection*, vol. 143, p. 105536, 2021.

[136] C. D. Mayee, G. T. Gujar, S. Dass, P. Balasubramanian, Y. Kapoor, and B. Choudhary, "In retrospect: managing an invasive pest, fall armyworm, *Spodoptera frugiperda*, in maize in India through digital and conventional networking pays off rich dividends towards crop sustainability," *Journal of plant diseases and protection (2006)*, vol. 128, no. 1, pp. 67–71, 2021.

[137] J. Y. Lee, S. Wang, A. J. Figueroa, R. Strey, D. B. Lobell, R. L. Naylor, and S. M. Gorelick, "Mapping Sugarcane in Central India with Smartphone Crowdsourcing," *Remote Sensing*, vol. 14, no. 3, 2022. [Online]. Available: https://www.mdpi.com/2072-4292/14/3/703

[138] L. C. Bertens, B. D. Broekhuizen, C. A. Naaktgeboren, F. H. Rutten, A. W. Hoes, Y. van Mourik, K. G. Moons, and J. B. Reitsma, "Use of expert panels to define the reference standard in diagnostic research: a systematic review of published methods and reporting," *PLoS medicine*, vol. 10, no. 10, p. e1001531, 2013.

[139] A. L. Baughman, K. M. Bisgard, M. M. Cortese, W. W. Thompson, G. N. Sanden, and P. M. Strebel, "Utility of composite reference standards and latent class analysis in evaluating the clinical accuracy of diagnostic tests for pertussis," *Clinical and Vaccine Immunology*, vol. 15, no. 1, pp. 106–114, 2008.

[140] P. D. McNicholas, "Model-based classification using latent gaussian mixture models," *Journal of Statistical Planning and Inference*, vol. 140, no. 5, pp. 1175–1181, 2010.

[141] C. Kalleshwaraswamy, R. Asokan, H. M. Swamy, M. Maruthi, H. Pavithra, K. Hegbe, S. Navi, S. Prabhu, G. E. Goergen *et al.*, "First report of the fall armyworm, *Spodoptera frugiperda* (JE Smith)(Lepidoptera: Noctuidae), an alien invasive pest on maize in India," 2018.

[142] S. R. Babu, R. Kalyan, S. Joshi, C. Balai, M. Mahla, and P. Rokadia, "Report of an exotic invasive pest the fall armyworm, *Spodoptera frugiperda* (J.E. Smith) on maize in southern rajasthan," *Journal of Entomology and Zoology Studies*, vol. 7, no. 3, pp. 1296–1300, 2019.

[143] R. N. Nagoshi, I. Dhanani, R. Asokan, H. Mahadevaswamy, C. M. Kalleshwaraswamy, and R. L. Meagher, "Genetic characterization of fall armyworm infesting South Africa and India indicate recent introduction from a common source population," *PLoS One*, vol. 14, no. 5, p. e0217755, 2019.

[144] C. Mallapur, A. K. Naik, S. Hagari, S. Prabhu, and R. Patil, "Status of alien pest fall armyworm, *Spodoptera frugiperda* (JE Smith) on maize in Northern Karnataka," *Journal of Entomology and Zoology Studies*, vol. 6, no. 6, pp. 432–436, 2018.

[145] D. Firake, G. Behere, S. Babu, and N. Prakash, "Fall armyworm: Diagnosis and management," *An Extension Pocket Book. Umiam-793*, vol. 103, 2019.

[146] J. M. Bland and D. G. Altman, "The odds ratio," *Bmj*, vol. 320, no. 7247, p. 1468, 2000.

[147] T. Benaglia, D. Chauveau, D. R. Hunter, and D. S. Young, "mixtools: An R Package for Analyzing Finite Mixture Models," *Journal of Statistical Software*, vol. 32, no. 6, pp. 1–29, 2009. [Online]. Available: http://www.jstatsoft.org/v32/i06/

[148] A. Zeimbekakis, E. D. Schifano, and J. Yan, "On Misuses of the Kolmogorov–Smirnov Test for One-Sample Goodness-of-Fit," *The American Statistician*, vol. 78, no. 4, pp. 481–487, 2024, published online: June 12, 2024.

[149] D. A. Grimes and K. F. Schulz, "Making sense of odds and odds ratios," *Obstetrics & Gynecology*, vol. 111, no. 2, pp. 423–426, 2008.

[150] S. Anandhi, V. Saminathan, P. Yasodha, S. S. J. Roseleen, P. Sharavanan, and V. Rajanbabu, "Correlation of fall armyworm *Spodoptera frugiperda* (J.E. Smith) with weather parameters in maize ecosystem," *Int. J. Curr. Microbiol. App. Sci*, vol. 9, no. 8, pp. 1213–1218, 2020.

[151] J. Menke, "Bivariate random-effects meta-analysis of sensitivity and specificity with the Bayesian SAS PROC MCMC: methodology and empirical evaluation in 50 meta-analyses," *Medical Decision Making*, vol. 33, no. 5, pp. 692–701, 2013.

[152] H. Srinivas, *New Development Patterns in India*. Dordrecht: Springer Netherlands, 2001, pp. 215–237. [Online]. Available: https://doi.org/10.1007/978-94-015-9670-$1_1$0

[153] A. Abbi, "Languages of India and India as a Linguistic Area," 2012. [Online]. Available: http://www.andamanese.net/LanguagesofIndiaandIndiaasalinguisticarea.pdf

[154] P. Schlattmann, "Tutorial: statistical methods for the meta-analysis of diagnostic test accuracy studies," *Clinical Chemistry and Laboratory Medicine (CCLM)*, vol. 61, no. 5, pp. 777–794, 2023. [Online]. Available: https://doi.org/10.1515/cclm-2022-1256

[155] J. Lee, K. W. Kim, S. H. Choi, J. Huh, and S. H. Park, "Systematic review and meta-analysis of studies evaluating diagnostic test accuracy: a practical review for clinical researchers-part II. Statistical methods of meta-analysis," *Korean journal of radiology*, vol. 16, no. 6, pp. 1188–1196, 2015.

[156] Y. Tsekhmister, "Effectiveness of case-based learning in medical and pharmacy education: A meta-analysis," pp. em515–em515, 2023.

[157] Y. Wu, B. Levis, K. E. Riehm, N. Saadat, A. W. Levis, M. Azar, D. B. Rice, J. Boruff, P. Cuijpers, S. Gilbody *et al.*, "Equivalency of the diagnostic accuracy of the PHQ-8 and PHQ-9: a systematic review and individual participant data meta-analysis," *Psychological medicine*, vol. 50, no. 8, pp. 1368–1380, 2020.

[158] C. Xu, L. Furuya-Kanamori, J. S. W. Kwong, S. Li, Y. Liu, and S. A. R. Doi, "Methodological issues of systematic reviews and meta-analyses in the field of sleep medicine: A meta-epidemiological study." *Sleep Medicine Reviews*, vol. 57, 2021.

[159] M. N. Plana, I. Arevalo-Rodriguez, S. Fernández-García, J. Soto, M. Fabregate, T. Pérez, M. Roqué, and J. Zamora, "Meta-DiSc 2.0: a web application for meta-analysis of diagnostic test accuracy data," *BMC Medical Research Methodology*, vol. 22, no. 1, p. 306, 2022.

[160] J. B. Reitsma, A. S. Glas, A. W. S. Rutjes, R. J. P. M. Scholten, P. M. Bossuyt, and A. H. Zwinderman, "Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews," *Journal of Clinical Epidemiology*, vol. 58, no. 10, pp. 982–990, 2005. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0895435605001629

[161] V. N. Nyaga and M. Arbyn, "Metadata: a Stata command for meta-analysis and meta-regression of diagnostic test accuracy data â a tutorial," *Archives of Public Health*, vol. 80, no. 1, p. 95, 2022.

[162] J. Pambabay-Calero, S. Bauz-Olvera, A. Nieto-Librero, A. Sánchez-García, and P. Galindo-Villardón, "Meta-analysis of diagnostic accuracy studies with multiple thresholds: Comparison of different approaches," *Biometrical Journal*, vol. 62, no. 1, pp. 1–15, 2020.

[163] Y. Shi, H. Fan, L. Li, Y. Hou, F. Qian, M. Zhuang, B. Miao, and S. Fei, "The value of machine learning approaches in the diagnosis of early gastric cancer: a systematic review and meta-analysis," *World J Surg Oncol.*, vol. 22, p. 40, 2024.

[164] M. Flor, M. Weiß, T. Selhorst, C. Müller-Graf, and M. Greiner, "Validation of a generic Bayesian method for prevalence estimation under misclassification," 2020.

[165] M. D. Teare, B. C. Author, and C. D. Last, "Sample size requirements to estimate key design parameters from external pilot randomised controlled trials:

a simulation study," *Trials*, vol. 15, no. 1, p. 264, 2014. [Online]. Available: http://www.trialsjournal.com/content/15/1/264

[166] M. Borenstein, L. V. Hedges, J. P. T. Higgins, and H. R. Rothstein, "A basic introduction to fixed-effect and random-effects models for meta-analysis," *Research synthesis methods*, vol. 1, no. 2, pp. 97–111, 2010.

[167] R. D. Riley, I. Ahmed, T. P. A. Debray, B. H. Willis, J. P. Noordzij, J. P. T. Higgins, and J. J. Deeks, "Summarising and validating test accuracy results across multiple studies for use in clinical practice," *Statistics in medicine*, vol. 34, no. 13, pp. 2081–2103, 2015.

[168] Z. Zhang, "A note on wishart and inverse wishart priors for covariance matrix," *Journal of Behavioral Data Science*, vol. 1, no. 2, pp. 119–126, 2021.

[169] D. J. Lunn, A. Thomas, N. Best, and et al., "WinBUGS - A Bayesian modelling framework: Concepts, structure, and extensibility," *Statistics and Computing*, vol. 10, pp. 325–337, 2000.

[170] S. Sturtz, U. Ligges, and A. Gelman, "R2WinBUGS: A Package for Running WinBUGS from R," *Journal of Statistical Software*, vol. 12, no. 3, pp. 1–16, 2005.

[171] T. F. Monaghan, S. N. Rahman, C. W. Agudelo, A. J. Wein, J. M. Lazar, K. Everaert, and R. R. Dmochowski, "Foundational statistical principles in medical research: sensitivity, specificity, positive predictive value, and negative predictive value," *Medicina*, vol. 57, no. 5, p. 503, 2021.

[172] M. H. Murad, L. Lin, H. Chu, B. Hasan, R. A. Alsibai, A. S. Abbas, R. A. Mustafa, and Z. Wang, "The association of sensitivity and specificity with disease prevalence: analysis of 6909 studies of diagnostic test accuracy," *CMAJ*, vol. 195, no. 27, pp. E925–E931, 2023.

[173] N. Brown, A. Pérez-Sierra, P. Crow, and S. Parnell, "The role of passive surveillance and citizen science in plant health," *CABI Agriculture and Bioscience*, vol. 1, pp. 1–16, 2020.

[174] K. L. Kier, "Biostatistical applications in epidemiology," *Pharmacotherapy: The Journal of Human Pharmacology and Drug Therapy*, vol. 31, no. 1, pp. 9–22, 2011.

[175] N. Speybroeck, B. Devleesschauwer, L. Joseph, and D. Berkvens, "Misclassification errors in prevalence estimation: Bayesian handling with care," *International journal of public health*, vol. 58, pp. 791–795, 2013.

[176] P. J. Diggle, "Estimating prevalence using an imperfect test," *Epidemiology Research International*, vol. 2011, 2011.

[177] W. J. Rogan and B. Gladen, "Estimating prevalence from the results of a screening test," *American journal of epidemiology*, vol. 107, no. 1, pp. 71–76, 1978.

[178] F. I. Lewis and P. R. Torgerson, "A tutorial in estimating the prevalence of disease in humans and animals in the absence of a gold standard diagnostic," *Emerging themes in epidemiology*, vol. 9, pp. 1–8, 2012.

[179] T. Dhar, S. Bhattacharya, H. Chatterjee, S. Senapati, P. Bhattacharya, P. Poddar, T. Ashika, and T. Venkatesan, "Occurrence of fall armyworm *Spodoptera frugiperda* (JE Smith)(Lepidoptera: Noctuidae) on maize in West Bengal, India and its field life table studies," *Journal of Entomology and Zoology Studies*, vol. 7, no. 4, pp. 869–875, 2019.

[180] S. K. Sharma, "First report of fall armyworm, *Spodoptera frugiperda* (JE Smith)(Lepidoptera: Noctuidae) incidence in Himachal Pradesh (HP), India," *Journal of Entomological Research*, vol. 45, no. 1, pp. 159–164, 2021.

[181] J. E. Fischer, L. M. Bachmann, and R. Jaeschke, "A readers' guide to the interpretation of diagnostic test properties: clinical example of sepsis," *Intensive care medicine*, vol. 29, pp. 1043–1051, 2003.

[182] M. Flor, M. Weiß, T. Selhorst, C. Müller-Graf, and M. Greiner, "Comparison of Bayesian and frequentist methods for prevalence estimation under misclassification," *BMC Public Health*, vol. 20, pp. 1–10, 2020.

[183] N. Speybroeck, N. Praet, F. Claes, N. Van Hong, K. Torres, S. Mao, P. Van den Eede, T. Thi Thinh, D. Gamboa, T. Sochantha *et al.*, "True versus apparent malaria infection prevalence: the contribution of a Bayesian approach," *PloS one*, vol. 6, no. 2, p. e16705, 2011.

[184] I. Sahlu and A. B. Whittaker, "Obtaining Prevalence Estimates of Coronavirus Disease 2019: A Model to Inform Decision-Making," *American Journal of*

Epidemiology, vol. 190, no. 8, pp. 1681–1688, 04 2021. [Online]. Available: https://doi.org/10.1093/aje/kwab079

[185] F. Habibzadeh, P. Habibzadeh, and M. Yadollahie, "The apparent prevalence, the true prevalence," *Biochem Med (Zagreb)*, vol. 32, no. 2, p. 020101, 2022.

[186] B. Carpenter, A. Gelman, M. D. Hoffman, D. Lee, B. Goodrich, M. Betancourt, M. A. Brubaker, J. Guo, P. Li, and A. Riddell, "Stan: A probabilistic programming language," *Journal of statistical software*, vol. 76, 2017.

[187] D. B. Sisodiya, B. L. Raghunandan, N. A. Bhatt, H. S. Verma, C. P. Shewale, B. G. Timbadiya, and P. K. Borad, "The fall armyworm, *Spodoptera frugiperda* (JE Smith)(Lepidoptera: Noctuidae); first report of new invasive pest in maize fields of Gujarat, India," *Journal of Entomology and Zoology Studies*, vol. 6, no. 5, pp. 2089–2091, 2018.

[188] A. Chormule, N. Shejawal, C. M. Sharanabasappa, R. Asokan, H. M. Swamy, and Z. Studies, "First report of the fall Armyworm, *Spodoptera frugiperda* (JE Smith)(Lepidoptera, Noctuidae) on sugarcane and other crops from Maharashtra, India," *Journal of entomology and zoology studies*, vol. 7, no. 1, pp. 114–117, 2019.

[189] G. Shankar and Y. Adachi, "First report of the occurrence of fall armyworm, *Spodoptera frugiperda* (JE Smith)(Lepidoptera: Noctuidae) on Ginger (Zingiber officinale) in Haveri district, Karnataka, India," *J. Entomol. Zool. Stud*, vol. 7, no. 5, pp. 78–80, 2019.

[190] U. Venkateswarlu, M. Johnson, R. Narasimhulu, and T. Muralikrishna, "Occurrence of the fall armyworm, *Spodoptera frugiperda* (JE Smith)(Lepidoptera, Noctuidae), a new pest on bajra and sorghum in the fields of agricultural research station, Ananthapuramu, Andhra Pradesh, India," *Journal of Entomology and Zoology Studies*, vol. 6, no. 6, pp. 811–813, 2018.

[191] S. Deole and N. Paul, "First report of fall army worm, *Spodoptera frugiperda* (JE Smith), their nature of damage and biology on maize crop at Raipur, Chhattisgarh," *Journal of Entomology and Zoology Studies*, vol. 6, no. 6, pp. 219–221, 2018.

[192] J. Srikanth, N. Geetha, B. Singaravelu, T. Rajasubramaniam, P. Mahesh, and e. a. Saravanan, "First report of occurrence of fall armyworm, *Spodoptera frugiperda* in sugarcane

from Tamil Nadu, India," *Journal of Sugarcane Research*, vol. 8, no. 2, pp. 195–202, 2018.

[193] R. Vishwakarma, K. Pragya, S. Patidar, S. B. Das, and A. Nema, "First report of fall army worm, *Spodoptera frugiperda* (JE Smith)(Lepidoptera: Noctuidae) on maize (*Zea mays*) from Madhya Pradesh, India," *Journal of Entomology and Zoology Studies*, vol. 8, no. 6, pp. 819–823, 2020.

[194] A. Kumar, S. Mehra, M. singh Jaglan, S. Yadav, S. singh Yadav *et al.*, "Life History Parameters of Fall Armyworm, *Spodoptera frugiperda* (JE Smith) on Maize in Haryana State," *Annals of Arid Zone*, vol. 63, no. 1, pp. 61–65, 2024.

[195] H. F. Nahrung and A. J. Carnegie, "Non-native Forest Insects and Pathogens in Australia: Establishment, Spread, and Impact," *Frontiers in Forests and Global Change*, vol. 3, 2020. [Online]. Available: https://www.frontiersin.org/articles/10.3389/ffgc.2020.00037

[196] J. E. Aukema, D. G. McCullough, B. Von Holle, A. M. Liebhold, K. Britton, and S. Frankel, "Historical accumulation of nonindigenous forest pests in the continental United States," *BioScience*, vol. 60, no. 11, pp. 886–897, 2010.

[197] J. C. Davis, J. P. Shannon, M. J. Van Grinsven, N. W. Bolton, J. W. Wagenbrenner, R. K. Kolka, and T. G. Pypker, "Nitrogen cycling responses to simulated emerald ash borer infestation in Fraxinus nigra-dominated wetlands," *Biogeochemistry*, vol. 145, no. 3, pp. 275–294, 2019.

[198] E. G. Brockerhoff and A. M. Liebhold, "Ecology of forest insect invasions," *Biological Invasions*, vol. 19, no. 11, pp. 3141–3159, 2017.

[199] E. J. Rebek, D. A. Herms, and D. R. Smitley, "Interspecific variation in resistance to emerald ash borer (Coleoptera: Buprestidae) among North American and Asian ash (*Fraxinus spp.*)," *Environmental entomology*, vol. 37, no. 1, pp. 242–246, 2014. [Online]. Available: https://doi.org/10.1603/0046-225X(2008)37[242:IVIRTE]2.0.CO;2

[200] C. Villari, D. A. Herms, J. G. A. Whitehill, D. Cipollini, and P. Bonello, "Progress and gaps in understanding mechanisms of ash tree resistance to emerald ash borer, a model for wood-boring insects that kill angiosperms," *New Phytologist*, vol. 209, no. 1, pp. 63–79, 2016.

[201] S. Fleck, C. Tomlin, F. da Silva Coelho, M. Richter, E. Danielsen, N. Backenstose, T. Krabbenhoft, C. Lindqvist, and V. Albert, "High quality long-read genomes produced from single MinION flow cells clarify polyploid and demographic histories of critically endangered ash species (*Fraxinus*: Oleaceae)," 2022, pREPRINT (Version 1) available at Research Square [https://doi.org/10.21203/rs.3.rs-2350866/v1].

[202] IUCN, "The IUCN Red List of Threatened Species. Version 2022-2," https://www.iucnredlist.org/, 2023, accessed 5 Jan 2023.

[203] Y. Dang, K. Wei, X. Wang, J. J. Duan, D. E. Jennings, and T. M. Poland, "Introduced plants induce outbreaks of a native pest and facilitate invasion in the plants' native range: Evidence from the emerald ash borer," *Journal of Ecology*, vol. 110, no. 3, pp. 593–604, 2021.

[204] T. Alamo, D. G. Reina, P. MillÃ¡n Gata, V. M. Preciado, and G. Giordano, "Data-driven methods for present and future pandemics: Monitoring, modelling and managing," *Annual Reviews in Control*, vol. 52, pp. 448–464, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1367578821000419

[205] J. M. Marshall, A. J. Storer, I. Fraser, J. A. Beachy, and V. C. Mastro, "Effectiveness Of Differing Trap Types For the Detection Of Emerald Ash Borer (Coleoptera: Buprestidae)," *Environ Entomol*, vol. 38, pp. 1226–1234, 2009.

[206] S. G. Jenkins, P. G. Oduor, L. Kotchman, and M. Kangas, "A Cohort Model For Ash Mortality Risk Due To Potential Emerald Ash Borer Infestation," *Biophysics*, vol. 06, no. 2, pp. 38–51, 2016.

[207] M. A. Lewis, *Spatial ecology: The Role of Space in Population Dynamics and Interspecific Interactions (MPB-30).* Princeton: Princeton University Press, 1997, ch. Variability, patchiness, and jump dispersal in the spread of an invading population, pp. 46–69.

[208] A. T. Peterson, "Predicting the geography of speciesâ invasions via ecological niche modeling," *The quarterly review of biology*, vol. 78, no. 4, pp. 419–433, 2003.

[209] M. S. Lau, G. J. Gibson, H. Adrakey, A. McClelland, S. Riley, J. Zelner, G. Streftaris, S. Funk, J. Metcalf, B. D. Dalziel *et al.*, "A mechanistic spatio-temporal framework for modelling individual-to-individual transmissionâwith an application to the 2014-2015

west africa ebola outbreak," *PLoS computational biology*, vol. 13, no. 10, p. e1005798, 2017.

[210] P. PyšEK and P. E. Hulme, "Spatio-temporal dynamics of plant invasions: linking pattern to process," *Ecoscience*, vol. 12, no. 3, pp. 302–315, 2005.

[211] S. F. Ward, S. Fei, and A. M. Liebhold, "Temporal dynamics and drivers of landscape-level spread by emerald ash borer," *Journal of Applied Ecology*, vol. 57, no. 6, pp. 1020–1030, 2020.

[212] B. A. Jones, "Can invasive species lead to sedentary behavior? the time use and obesity impacts of a forest-attacking pest," *Journal of Environmental Economics and Management*, vol. 119, p. 102800, 2023.

[213] P. Bettinger, K. Boston, J. P. Siry, and D. L. Grebner, "Chapter 2 - Valuing and Characterizing Forest Conditions," in *Forest Management and Planning (Second Edition)*, second edition ed., P. Bettinger, K. Boston, J. P. Siry, and D. L. Grebner, Eds. Academic Press, 2017, pp. 21–63. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9780128094761000023

[214] L. R. Iverson, A. Prasad, J. Bossenbroek, D. Sydnor, M. W. Schwartz *et al.*, "Modeling potential movements of the emerald ash borer: the model framework," *Advances in threat assessment and their application to forest and rangeland management. Portland, OR: US Department of Agriculture, Forest Service, Pacific Northwest and Southern Research Stations. p*, pp. 581–97, 2010.

[215] T. U. S. D. of Agriculture (USDA) Forest Service, "Live tree species basal area of the contiguous United States (2000-2009)," accessed on February 14, 2024. [Online]. Available: https://www.fs.usda.gov/rds/archive/catalog/RDS-2013-0013

[216] S. J. Fahrner, J. P. Lelito, and B. H. Aukema, "The influence of temperature on the flight capacity of emerald ash borer *Agrilus planipennis* and its parasitoid, *Tetrastichus planipennisi*: implications to biological control," *BioControl*, vol. 60, pp. 437–449, 2015.

[217] W. Paul and J. Baschnagel, *Stochastic processes, From Physics to Finance*. Springer, 1999.

[218] A. Kleczkowski, A. Hoyle, and P. McMenemy, "One model to rule them all? modelling approaches across onehealth for human, animal and plant epidemics," *Philosophical Transactions of the Royal Society B*, vol. 374, no. 1775, p. 20180255, 2019.

[219] R. A. Taylor, L. S. Bauer, T. M. Poland, and K. N. Windell, "Flight performance of *Agrilus planipennis* (Coleoptera: Buprestidae) on a flight mill and in free flight," *Journal of Insect Behavior*, vol. 23, no. 2, pp. 128–148, 2010.

[220] D. Yemshanov, R. G. Haight, F. H. Koch, B. Lu, R. Venette, D. B. Lyons, T. Scarr, and K. Ryall, "Optimal allocation of invasive species surveillance with the maximum expected coverage concept," *Diversity and Distributions*, vol. 21, no. 11, pp. 1349–1359, 2015.

[221] A. Baskar, "Simple single and multi-facility location models using great circle distance," vol. 37, p. 01001, 2021.

[222] T. Sing, O. Sander, N. Beerenwinkel, and T. Lengauer, "ROCR: visualizing classifier performance in R," *Bioinformatics*, vol. 21, no. 20, pp. 3940–3941, 08 2005. [Online]. Available: https://doi.org/10.1093/bioinformatics/bti623

[223] J. M. Marshall, M. A. Miller, J. P. Lelito, and A. J. Storer, "Latitudinal variation in body size of *Agrilus planipennis* and relationship with fecundity," *Agricultural and forest entomology*, vol. 15, no. 3, pp. 294–300, 2013.

[224] R. A. Haack and T. R. Petrice, "Mortality of bark-and wood-boring beetles (Coleoptera: Buprestidae, Cerambycidae, and Curculionidae) in naturally infested heat-treated ash, birch, oak, and pine bolts," *Journal of Economic Entomology*, vol. 115, no. 6, pp. 1964–1975, 2022.

[225] J. Fan, M. Haseeb, Q. Ren, T. Tian, R. Zhang, and P. Wu, "Factoring distribution and prevalence of Fall armyworm in southwest China," *Journal of Applied Entomology*, vol. 145, no. 4, pp. 295–302, 2021.

[226] C. E. Hausman, J. F. Jaeger, and O. J. Rocha, "Impacts of the emerald ash borer (EAB) eradication and tree mortality: potential for a secondary spread of invasive plant species," *Biological invasions*, vol. 12, pp. 2013–2023, 2010.

[227] K. S. Knight, J. P. Brown, and R. P. Long, "Factors affecting the survival of ash (*Fraxinus spp.*) trees infested by emerald ash borer (*Agrilus planipennis*)," *Biological Invasions*, vol. 15, pp. 371–383, 2013.