Journal of Applied Entomology

WILEY

JOURNAL OF APPLIED ENTOMOLOGY

REVIEW ARTICLE OPEN ACCESS

Spatial Interpolation in Applied Insect Ecology: A Review, Including Guidelines and Datasets for Practical Use

Janne Heusler[1] | Jonas Funk[1] | Andreas Wagner[1,2]

[1]Karlsruhe University of Applied Sciences, Karlsruhe, Germany | [2]Fraunhofer Institute for Industrial Mathematics ITWM, Kaiserslautern, Germany

Correspondence: Jonas Funk (jonas.funk@h-ka.de)

Received: 23 April 2025 | **Revised:** 23 April 2025 | **Accepted:** 26 May 2025

Funding: This work was supported by Carl-Zeiss-Stiftung.

Keywords: entomology | geostatistics | kriging | literature review | spatial interpolation | species distribution modelling

ABSTRACT

Spatial interpolation represents a fundamental approach in applied insect ecology, offering insight into species distributions and supporting biodiversity analysis, pest management and disease vector mapping. Insects—including important pollinators—face escalating threats due to habitat loss, climate change and anthropogenic pressures. As data-driven decisions become more critical in addressing these ecological challenges, spatial interpolation techniques such as kriging and regression-based models have become essential for estimating insect abundance in unsampled areas. This paper offers an in-depth review of both geostatistical and non-geostatistical methods employed in insect ecology, including ordinary kriging, universal kriging and machine learning-based methods such as random forests and maximum entropy. We present a structured overview of their applications in pest management, disease vector mapping and biodiversity monitoring, and we provide practical guidelines for selecting appropriate spatial interpolation methods. In addition, we present several datasets that can support case studies in spatial modelling for insect ecology. Our findings underscore the advantages of integrating geostatistical approaches with environmental variables to enhance the accuracy of species distribution models. This review serves as a resource for entomologists and researchers seeking to advance ecological monitoring and management through spatial interpolation techniques.

1 | Introduction

The role of insects and pollinators in maintaining healthy ecosystems is fundamental; however, they are facing unprecedented threats around the world. The decline of these species has significant implications for biodiversity and human food security, as many crops rely on pollination services (Gallai et al. 2009).

These threats include climate change, which alters habitats and disrupts life cycles, leading to mismatches in the timing of pollinator activities and plant flowering (Vanbergen and Insect Pollinators Initiative 2013). Habitat loss, primarily driven by agricultural expansion and urbanisation, leads to the fragmentation of natural environments, reducing available resources and nesting sites for pollinators (Potts et al. 2010). Moreover, human activities, particularly the use of pesticides and pollutants, reduce insect populations and biodiversity, further destabilising ecosystems (Goulson et al. 2015). These factors threaten individual species and undermine essential ecological services, such as pollination and decomposition, critical to ecosystem health. As ecosystems become increasingly fragile, the protection of insect populations becomes vital, requiring a comprehensive understanding of pollinator behaviour and the environmental changes they face (Potts et al. 2016).

Expanding efforts to collect and analyse data on these threats is crucial to making informed, data-driven decisions that can help reverse or mitigate declines. However, the prevailing monitoring techniques either rely on time-consuming field observations or the use of invasive traps. Thus, they are unable to provide efficient

---

https://doi.org/10.1111/jen.13455

monitoring and preserve insect species on extensive spatial and temporal scales (Montgomery et al. 2021; Klink et al. 2022).

This paper is part of the BeeVision project, which addresses these challenges by combining event cameras and artificial intelligence to provide a cost-effective and scalable solution to monitor pollinator populations. For example, the spatial distribution of insects can be mapped using such data. In the face of these challenges, an understanding of spatial patterns is increasingly important (Vinatier et al. 2011). To achieve this, spatial interpolation techniques, which employ known data points to estimate unknown values, are widely used in applied insect ecology to create maps from data to facilitate understanding and management of environmental processes (Sciarretta and Trematerra 2014). For example, to understand the distribution of disease vectors to improve prevention and monitoring measures (Eisen and Eisen 2011), to map the distribution of pests in crops to combat infestations (Liebhold et al. 2003; Moral García 2006) or to analyse biodiversity to gain additional insights (Ferrier 2002). This paper contributes to this broader context by offering a comprehensive review and practical guidance on the application of spatial interpolation methods in entomology.

The paper makes three contributions:

1. *Literature review*: It provides a structured and comprehensive review of existing research on spatial interpolation in entomology, focusing on both geostatistical and non-geostatistical methods. The literature is structured after the three main applications: pest management, mapping of disease vectors and analysis of biodiversity.
2. *Guidelines for researchers*: This work offers practical guidelines on selecting appropriate spatial interpolation methods for different entomological applications. It outlines criteria for data collection, applicability, the interpolation areas, model evaluation and software availability, thereby supporting researchers in making informed methodological choices.
3. *Datasets for case studies*: The paper presents a set of available datasets that can be used to conduct case studies on pest management, biodiversity analysis and disease vector mapping. This section aims to facilitate further research by providing valuable resources for testing and validation of spatial interpolation models.

The relevance of this work lies in its potential to assist both entomologists and researchers in related fields involved in ecological monitoring and pest management. By synthesising and categorising existing methods, offering clear guidelines and providing datasets, this work serves as an important resource and aims to enhance the capacity of researchers to monitor and analyse insect populations more effectively. The paper is structured as follows: Firstly, some fundamental principles of spatial interpolation are introduced. Secondly, the existing literature on spatial interpolation in entomology is reviewed, with a particular focus on methods and applications. Thirdly, practical guidelines are provided to assist researchers in selecting suitable interpolation techniques based on specific needs. Finally, a selection of datasets is presented, accompanied by recommendations for their use in future research.

2 | Materials and Methods

2.1 | Spatial Interpolation

This chapter introduces spatial interpolation, emphasising its importance in environmental science for predicting values at unsampled locations. It introduces the various interpolation methods used in the existing literature, providing a theoretical basis for the paper and justifying the focus on kriging and other geostatistical approaches in the context of insect ecology. It discusses the mathematical principles behind interpolation methods and how kriging uses spatial autocorrelation to make predictions. The basic concept of spatial interpolation is to utilise known data points to predict unknown values based on a given function (Li and Heap 2014).

Table 1 shows the different spatial interpolation methods (SIM) applied in the literature on entomology reviewed in this paper, categorised in three distinct groups. Methods, namely kriging, are based on the principles of geostatistics, methods outside the field of geostatistics, and methods combining approaches from both groups. In total, there are 15 different methods covered which have been applied in existing literature, with 66.6% being some variation of kriging. It is important to note that there are more existing methods to interpolate data than the ones categorised in Table 1. Many of these methods have been used in studies outside the field of entomology. In this review, only interpolation methods which have been used or cited in existing studies in the field of entomology are categorised. Estimations of most SIMs can be represented as weighted averages of sampled data. They all share the same general estimation equation, as seen in Equation (1): (Li 2008).

$$\hat{z}(x_0) = \sum_{i=1}^{n} \lambda_i z(x_i) \tag{1}$$

In this case $\hat{z}$ is the predicted value at the point x_0, z is the observed value at the point x_i, λ_i is the weight assigned to the

TABLE 1 | Overview of spatial interpolation methods.

Geostatistics	Non-geostatistics	Combined
Ordinary kriging	Inverse distance weight	Regression kriging
Universal kriging	Generalised linear models	Regression space time simple kriging
Co-kriging	Regression tree	
Indicator kriging	Random forest	
Block kriging	Generalised additive models	
Multi Gaussian co-kriging	Max entropy	
Binomial areal kriging		
Space time ordinary kriging		

sample point by the model and n is the number of total sample points. The weights are designed to sum up to 1 (Webster and Oliver 2007).

Geostatistics are a family of statistics that describe correlation through space to provide unbiased predictions of unknown values and a measure of prediction uncertainty (Oliver and Webster 2015). Positive spatial autocorrelation determines that observations in close spatial proximity are more likely to exhibit similar values than those at greater distances. In the field of geostatistics, this concept is key for modelling and predicting spatial patterns of geological and environmental variables (Isaaks and Srivastava 1989). When applying kriging, spatial autocorrelation is quantified using the semivariogram, a fundamental tool in geostatistics. Although technically referring to the semivariance, the term 'variogram' is often used interchangeably with 'semivariogram' in many geostatistical applications. The semivariogram measures the variance between sample values as a function of distance, providing insight into the spatial continuity and structure of the data (Cressie 2015). The semivariogram equation calculates the semivariance, which is half the average-squared difference between pairs of sample points against their distance.

$$\hat{y}(h) = \frac{1}{2m(h)} \sum_{i=1}^{m(h)} \left\{ z(x_i) - z(x_i + h) \right\}^2 \tag{2}$$

$\hat{y}(h)$ is the calculated semivariance for a specific distance interval. $z(x_i)$ and $z(x_i + h)$ are the sample values at two locations x_i and $x_i + h$ with the distance h. $m(h)$ is the number of sample pairs with the distance h, also called a lag.

All the key features of the semivariogram can be seen in Figure 1. The dots represent the actual calculated values of the semivariance at different distance intervals (lags). This is a plot of the empirical data showing how the semivariance changes with increasing distance between points (Gringarten and Deutsch 2001). Lags refer to the intervals between data points used to calculate the semivariogram. Each dot in the experimental semivariogram represents the average difference in values between points that are a certain distance (lag) apart. The semivariogram is computed by aggregating these differences over the various lag distances (Benndorf 2023). The nugget represents the semivariance at a distance of zero, which indicates a measurement error or microscale variations. A high nugget implies that the data has no clear spatial autocorrelation (Goulard and Voltz 1992). The sill represents the plateau at which the semivariogram stabilises, which indicates the total variance of the data. Once the sill is reached, no further correlation between data points is observed as distance increases and the semivariogram can no longer make any evaluation of the variance (Bohling 2005). The range marks the distance when the semivariogram reaches the sill. The range of the semivariogram represents the maximum distance at which autocorrelation can be observed in the dataset (Wackernagel 1995). The experimental semivariogram is a direct calculation of how variance between points changes with distance. To predict the spatial structure of the data and perform kriging or other spatial interpolations, mathematical models such as the Gaussian, Spherical, or Exponential semivariograms are fitted to the experimental data. The choice of model affects the interpolation, with each model providing a different assumption about how variance behaves as distance increases (Oliver and Webster 2015).

Kriging leverages the estimations by the semivariogram to determine weights and make predictions for unsampled locations (Cressie and Hawkins 1980). The core equation for the kriging estimator is similar to Equation (1):

$$\hat{Z}(x_0) - \mu = \sum_{i=1}^{n} \lambda_i \left[Z(x_i) = \mu(x_0) \right] \tag{3}$$

λ_i are the kriging weights estimated by minimising the variance, n is the number of samples and $\mu(x_0)$ is the mean of samples within the observed area. Some kriging models assume a known constant mean of the values across the sample area, which is

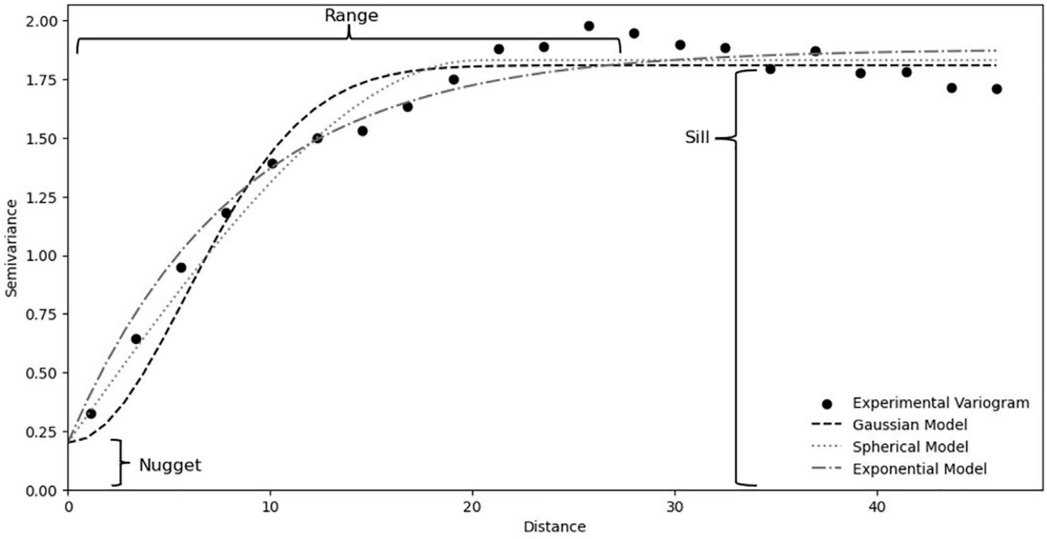

FIGURE 1 | Example of an experimental semivariogram with different models.

represented by μ (Wackernagel 1995). Similar to Equation (1), kriging uses the known data points around the prediction location x_0 along with their respective weights to estimate $\hat{Z}$. Next to making predictions, an important trait of kriging is its ability to calculate the prediction error of the model. The weights in Equation (3) are calculated with the goal of minimising that error. To do that, kriging applies two conditions: unbiasedness to the data, meaning that the predictions are not systemically biased but represent an average across all sample points, and minimisation of prediction variance, meaning that the estimated values should be as close as possible to the true values (Benndorf 2023). This makes kriging suitable for mapping variables like soil properties (Voltz and Webster 1990; Kravchenko and Bullock 1999), precipitation (Di Piazza et al. 2011) and pollution concentrations (Janssen et al. 2008; Wong et al. 2004), in which spatial correlation is significant (Goovaerts 1997). Kriging can be adapted for various forms of data through its variants, allowing it to handle both isotropic and anisotropic data, univariate and multivariate data and continuous and binary data. Anisotropic data has different properties depending on the direction in which it is measured. This means the data exhibits different behaviour or characteristics along different axes or directions. Isotropic data has uniform properties in all directions (Callister and Rethwisch 2018). Univariate data involves a single variable, while multivariate data involves multiple variables at once, which can be analysed for relationships, interactions and patterns across these variables (Montgomery and Runger 2018). Continuous data can take any value within a range, while binary data consists of only two possible values, often representing categories or states like true or false (Agresti 2019). The most common types of kriging are ordinary kriging (OK), universal kriging (UK), incorporating a trend function and co-kriging (CoK), which incorporates secondary variables into the calculations.

Besides spatial interpolation based on geostatistics, there are other methods which do not use information about the location to make predictions. Non-geostatistical interpolation methods may be more appropriate in scenarios where environmental variables are more dependent on external factors such as altitude, land cover, or climate conditions than on spatial location alone, or where sampling of the primary variable is more difficult than sampling of the external factors (Dale and Fortin 2014). These methods, as seen in Table 1, in literature often referred to as species distribution models (SDM) or ecological niche models, do not inherently model spatial correlation but can integrate environmental predictors to enhance interpolation accuracy. Applying these methods involves leveraging their capacity to integrate diverse data sources, which is important in dynamic environmental systems (Li 2008). Inverse distance weighting (IDW) estimates unknown values based on known values from surrounding locations. As with kriging, it assumes that points closer to the unsampled location have more impact on the predicted value than those further away. The core equation for IDW weights is:

$$\lambda_i = \frac{\frac{1}{d_i^p}}{\sum_{i=1}^{n} \frac{1}{d_i^p}} \tag{4}$$

d_i is the distance between sampled and unknown locations, n is the number of samples and p is a power parameter, which determines how fast the impact of sample points decreases. Weights decrease as distance increases (Shepard 1968). A higher value of p increases the weight of nearby samples. IDW performs well with regularly spaced data and can be used when sampling data is scarce. Unlike kriging, it is not able to cluster samples. Due to its simplicity, it is often used as a benchmark for kriging interpolation (Isaaks and Srivastava 1989). Several models are using regressions to make interpolations. Regressions explore a possible relationship between primary and secondary variables to make predictions (Draper and Smith 1998). Linear regression assumes a linear relationship between the variables and tries to find the optimal linear function to model this relationship (Hosmer Jr. et al. 2013) generalised linear models (GLM) extend regression by allowing the variables to have different distributions, accommodating various data types such as binary, count, or continuous. Based on the chosen function, GLMs apply different regression models (McCullagh 2019). GLMs themselves can be extended by generalised additive models (GAM) by allowing non-linear relationships between predictors and variables, which can improve results when modelling complex patterns in the data (Hastie and Tibshirani 1990). Random forests (RFs) are a machine learning technique that generates multiple decision trees trained on different subsets of the data that each produce a distinct output. RFs can perform both classification and regression, with the result being the aggregation of all trees to reduce training errors and enhance prediction accuracy (Breiman 2001). A similar method to RFs used for interpolation is regression trees (RTs). Opposed to RFs, they only create a single decision tree composed of decision nodes and leaf nodes. Prediction results are based on the mean value of the target variable in each leaf node, instead of an aggregation of different trees in the case of RFs (Breiman et al. 2017) maximum entropy (MaxEnt) is another machine learning method for modelling species distributions. It estimates the probability distribution of a species across a landscape by finding the probability distribution (i.e., the most widespread or uniform distribution) subject to constraints derived from environmental variables at known locations of presence (Elith et al. 2011).

To accompany both groups, some approaches combine geostatistics and non-geostatistics to use and handle external data while still taking spatial correlation into account. Many of these models are combining regression models with kriging. They first model the spatial trend using the regression and then apply kriging on the residuals to capture the spatial autocorrelation to enhance the prediction accuracy (De Cesare et al. 2002). This method is particularly useful when the spatial trend is influenced by external variables. An additional advantage is the fact that it allows separate interpretation of both components, regression and kriging (Hengl 2007).

2.2 | Existing Work

This section presents an overview of existing literature on the application of spatial interpolation in the field of entomology. There is not a lot of related work which either focuses on environmental sciences as a whole (Li 2008), is no longer current (Liebhold et al. 2003), or only covers singular applications (Sciarretta and Trematerra 2014). Particular attention is paid to the application of methods such as kriging and its variants in pest management, biodiversity analysis and disease vector mapping. This section focuses on the different applications SIMs are used for, describing the scenarios they have been applied in and their results. Existing work on insect species distribution modelling can be grouped into three main applications: pest management, mapping of disease vectors and mapping and analysis of biodiversity. Table 2 shows the different interpolation methods that have been applied in studies which are going to be covered in the following sections. Each section starts with a table, structuring the studies according to the methods used, the species observed, the sampling period and the secondary variables used for modelling, followed by a review of the results and insights of each study. The goal of the review is to present an overview of the variety of different interpolation methods that already have been used in research.

Pest management is represented the most in the literature and therefore differentiated between local applications and regional/global applications of pest management. Areas below a scale of $5\,\text{km}^2$ are considered local, and areas above count towards regional/global.

2.3 | Pest Management

In the following two sections, an overview of existing applications of SIMs for pest management on both a local and regional/global scale is provided. Pest management involves understanding the behaviour, life cycle and population dynamics of pest insects to develop effective control strategies. Spatial interpolation methods have emerged as useful tools in this domain, offering approaches for spatial analysis and prediction (Sciarretta and Trematerra 2014). It is noteworthy that some literature does not define the specific kriging method used and will only be referred to as kriging in this review.

2.3.1 | Local

Table 3 summarises the literature reviewed in this chapter, including the SIM used, the species observed, the sampling period, and secondary variables, where appropriate for the methods. The samples used in these studies can either come from samples collected during the study itself, or from other sources, such as institutions that conduct regular surveys of species occurrence. In the latter case, the sampling period may be much longer.

Weisz et al. (1995a) compared OK with four different IDW functions and found that OK produces the best predictions, followed by IDW functions with higher exponents. All interpolators showed weakness in areas with low mean densities. In an additional study on the Colorado potato beetle using kriging, Weisz et al. (1995b) concluded that making targeted interventions based on maps created by kriging could result in time savings and reduced pesticide load for the environment. But designing the sampling program and interpolation needs to be appropriate for the observed pest. Reay-Jones (2010) found that IDW is useful for interpolation of spatial patterns with scarce sampling data. Several studies saw a high variance in results using different semivariogram models. This outlines the importance of testing different models when applying spatial

TABLE 2 | Methodological approaches across different fields.

Method	Pest management		Disease vectors	Biodiversity
	Local	**Regional/global**		
Geostatistics				
Ordinary kriging	x	x	x	x
Universal kriging			x	x
Co-kriging		x	x	x
Other kriging methods	x	x		
Non-geostatistics				
IDW	x			x
GLM			x	x
RT			x	
RF			x	
GAM				x
MaxEnt		x	x	
Logistic regression		x		
Combined				
Regression kriging				
RSTK			x	

TABLE 3 | Summary of studies on local pest management.

Study	Methods	Species	Sampling time frame
Brandhorst-Hubbard et al. (2002)	Kriging	Several soil invertebrates	2 months
Bressan et al. (2010)	OK	Cixiid planthoppers (*Pentastiridius leporinus*)	2 years
Ellsbury et al. (1998)	Kriging	Western (*Diabrotica virgifera*) and northern (*Diabrotica barberi*) corn rootworm	5 months
Martins et al. (2018)	OK	Tomato leafminer (*Tuta absoluta*)	6 months
Moral García (2006)	OK	Cotton bollworm (*Helicoverpa armigera*)	2 months
Park and Obrycki (2004)	OK	Corn leaf aphids (Homoptera: Aphididae), lady beetles (Coleoptera: Coccinellidae)	3 months
Reay-Jones (2010)	IDW	Cereal leaf beetle (*Oulema melanopus*)	2 years
Rijal et al. (2014)	OK	Grape root borer (*Vitacea polistiformis*)	4 years
Sciarretta et al. (2001)	OK	Plum fuit moth (*Cydia funebrana*)	12 months
Weisz et al. (1995a)	OK, IDW	Colorado potato beetle (*Leptinotarsa decemlineata*)	6 months
Weisz et al. (1995b)	Kriging	Colorado potato beetle	6 months

interpolation (Ellsbury et al. 1998; Rijal et al. 2014). In some cases, certain models were not able to identify spatial correlation at all, while other models identified significant correlation (Park and Obrycki 2004). In addition, Ellsbury et al. (1998) saw that kriging can be beneficial for pest management by offering alternative approaches to sampling and management of insects' pests. Several studies found OK to be helpful in analysing pests, such as their movement in fields (Moral García 2006), origin and direction of outbreaks (Martins et al. 2018) or their life cycle and preferred host areas (Bressan et al. 2010). Farmers can use these insights to develop sampling plans, assess infestations and improve pest management. Kriging was helpful to map soil infestation by several insect species using acoustic systems as a non-destructive tool for sampling (Brandhorst-Hubbard et al. 2002). If there is an unusual trap count or suboptimal sample distribution, OK is capable of interpolating data for the entire area, thus allowing the use of additional methods that incorporate supplementary environmental data (Sciarretta et al. 2001).

2.3.2 | Regional/Global

Defoliation caused by gypsy moths is a well-researched problem. Several studies used CoK to predict defoliation, as seen in Table 4. Hohn et al. (1993) found, when using time as a third dimension, that results tended to lag 1 year behind actual values, attributed to the model's tendency to adjust heavily to the values of the preceding year. In addition, the same researchers had better results using a logistic regression than CoK (Gribko et al. 1995). In contrast, Liebhold et al. (1998) saw high model errors with logistic regression across different models, using egg mass density estimates of gypsy moths and distance to the infested front as variables. Next to time, other variables are helpful in explaining spatial correlation using different CoK models, like elevation, slope and outbreak frequency (Cocco et al. 2010; Castrignanò et al. 2012). Several studies acknowledge the potential of OK to improve pest management by providing early information on outbreaks and seasonal movements, which can be used to intensify monitoring zones, delimit treatment areas and optimise countermeasures and trap placement (Farias et al. 2004; Duarte et al. 2015; Koch and Smith 2008). Kriging methods like OK and indicator kriging (IK) are helpful in supporting other interpolation methods, like SDMs, by providing additional information on the insect pests (Sharov et al. 1995; Koch and Smith 2008). When using OK on large areas, the assumption of a constant mean (stationarity) can hurt predictions. To improve the prediction quality, the area can be divided into several ecoregions with similar environmental traits, and OK can be performed separately for each subregion (Koch and Smith 2008). Using kriging improves the quality of hazard maps of potential infestation and drastically reduces the time required to develop maps compared to correlograms, but the accuracy of maps is impacted by the chosen grid size for the calculation (Kemp et al. 1989). An SDM like, the MaxEnt algorithm can predict current infestation and help analyse habitat suitability of pests to improve management. However, when using remotely sensed data, quality and availability can become a problem, since it can be costly, and external conditions can impact the quality. Furthermore, fine-tuning models to improve accuracy is crucial. Using models without adjusting parameters will lead to incomplete predictions. Furthermore, reducing model complexity and incorporating biological knowledge about species can improve both model accuracy and the ability to transfer the model to new areas. Reducing model complexity means focusing on fewer environmental variables which appear to have the highest impact on describing the distribution of a specific species (Makori et al. 2017; Srivastava et al. 2021).

TABLE 4 | Summary of studies on regional/global pest management.

Study	Methods	Species	Sampling time frame	Secondary variables[a]
Castrignanò et al. (2012)	Multi-Gaussian co-kriging	Olive fruit fly (*Bactrocera oleae*)	6 months	Elevation
Cocco et al. (2010)	CoK	Gypsy moth (*Lymantria dispar*)	24 years	Elevation, slope, outbreak frequency
Duarte et al. (2015)	OK	Oriental fruit moth (*Grapholita molesta*)	3 years	—
Farias et al. (2004)	OK	Sharpshooters (*Xylella fastidiosa*)	9 months	—
Gribko et al. (1995)	Logistic regression	Gypsy moths	39 years	Time
Hohn et al. (1993)	CoK	Gypsy moths	39 years	Time
Kemp et al. (1989)	BK	Rangeland grasshoppers (Orthoptera: Acrididae)	1 year	—
Koch and Smith (2008)	OK, SDM	Redbay ambrosia beetles (*Xyleborus glabratus*)	Not specified	Temperature, moisture, precipitation
Liebhold et al. (1998)	Logistic regression	Gypsy moths	3 years	Egg mass density, trapped males, defoliation in previous year, distance to population front
Makori et al. (2017)	MaxEnt	Small (*Aethina tumida*) and large (*Oplostomus haroldi*) hive beetle, honeycomb moth (*Galleria mellonella*), varroa mite (*Varroa destructor*)	3 months	Biotic variables, topographical, temperature, precipitation
Sharov et al. (1995)	Indicator kriging	Gypsy moths	5 years	—
Srivastava et al. (2021)	MaxEnt	Asian gypsy moth, Asian longhorned beetle (*Anoplophora glabripennis*)	Not specified	19 bioclimatic variables

[a] If applicable for the respective methods.

2.4 | Mapping of Disease Vectors

Another application is the mapping of disease vectors to improve monitoring and understand transmission patterns. Mapping disease vectors is a fundamental step in understanding and managing vector-borne diseases. Disease vectors, including mosquitoes, ticks and flies, have a significant impact on the transmission of pathogens that cause diseases such as Malaria, Dengue, Zika and Chagas disease. Accurate mapping of these vectors makes it possible to predict disease outbreaks, implement control measures, and ultimately reduce the prevalence of vector-borne diseases (Eisen and Eisen 2011).

Table 5 shows the different studies using spatial interpolation for this application. UK has been used by several studies to map mosquito and egg-mass abundance. Based on the results, UK is useful for mapping egg-laying hotspots and improving monitoring and prevention by allowing researchers to make more targeted control efforts (Richards et al. 2006; Ryan et al. 2004). The study by Cano et al. (2007) showed that using OK and UK helps to analyse the distribution and dynamics of disease vectors, and both methods are useful tools to discriminate zones with high and low abundance and to predict potential transmission of sickness. While Duncombe et al. (2013) were able to use OK to identify high transmission zones in certain time frames, OK also failed to observe spatial correlation in 43 out of the total 57 observed months. The sampling method has a significant impact on interpolation results. Azil et al. (2014) showed that by using different traps, the data revealed significantly higher spatial correlation. Instead of using sticky ovitraps to trap breeding females, they used a trap designed to capture adult mosquitoes in general. In addition, results of predictions can be improved by using CoK instead of OK with the Premise Condition Index, a measure of house and property condition related to mosquito breeding potential, and total potential breeding containers as variables (Azil et al. 2014). Several other studies reported similar results and were able to improve predictions by using multivariate kriging models. Giordano et al. (2018) observed that using all available landscape variables: elevation and proximity to conservation areas, provincial parks, wetlands and population centres, did not improve the spatial prediction. The observed species exhibited spatial autocorrelation with only a few and

TABLE 5 | Summary of studies on disease vector mapping.

Study	Methods	Species	Sampling time frame	Secondary variables[a]
Azil et al. (2014)	OK, CoK	Mosquito (*Aedes aegypti*)	6 weeks	Premise condition index, potential breeding containers
Cano et al. (2007)	OK, UK	Tsetse fly (*Glossina palpalis palpalis*)	18 months	—
Cianci et al. (2015)	RF, GLM	Mosquito	Not specified	11 environmental variables
Duncombe et al. (2013)	OK	Mosquito	57 months	—
Espinosa, Weinberg, et al. (2016)	MaxEnt	Mosquito	5 years	23 environmental, climatic, demographic variables
Giordano et al. (2018)	Universal CoK, simple CoK	12 west Nile virus vectors (Diptera: Culicidae)	12 years	Elevation, proximity to conservation areas, parks, wetlands, population centres
Richards et al. (2006)	UK	Mosquito	2 years	—
Ryan et al. (2004)	UK	Four mosquito species	16 weeks	—
Sedda et al. (2010)	Space–time OK, regression Space–time Simple kriging	Tsetse fly	15 months	Habitat area, mean patch size, patch size standard deviation
Zheng et al. (2019)	RT	Mosquito	20 years	Average temperatures, monthly rainfall

[a] If applicable for the respective methods.

differing variables. Sedda et al. (2010) found that by using environmental variables, except just time, results improved by comparing space–time OK and regression space–time simple kriging (RSTSK). Seasonality and total forest area (habitat) emerged as the primary determining factors for tsetse abundance. Several different SDMs, RTs, MaxEnt, RFs, GLM, were used to model mosquito abundance with similar goals. For example, to monitor and predict changes in the seasonal movements of pests to adapt prevention programs, identify high-risk areas to make targeted control measures and to model environmental suitability (Zheng et al. 2019; Espinosa, Polop, et al. 2016; Espinosa, Weinberg, et al. 2016; Cianci et al. 2015). These models use many variables to perform spatial interpolation compared to kriging. Some studies use up to 23 different environmental, climatic and demographic variables (Espinosa, Weinberg, et al. 2016). Cianci et al. (2015) built different models using up to 11 variables and found that maps created with RFs outperformed those generated with GLMs.

2.5 | Biodiversity

Mapping the insect biodiversity is a vital endeavour in entomology, providing critical insights into ecosystem health, the effects of climate change and species conservation. It can help scientists understand patterns of species richness, identify biodiversity hotspots and monitor changes over time. This can improve conservation planning, prioritise areas for protection, sustainably manage habitats and restore degraded ecosystems (Ferrier 2002).

As seen in Table 6, mapping biodiversity includes many secondary variables and is not just based on spatial correlation. Common variables include temperature, habitat, rainfall and elevation. A study by Monfared et al. (2013) supports this. They concluded that, when analysing insect biodiversity, multivariate kriging models that incorporate environmental variables tend to outperform univariate models like UK or OK. Several studies use GLM to predict insect biodiversity across spatial areas. GLMs are suitable since they can incorporate a high number of variables into the predictions (Hawkins and DeVries 2009; Plascencia and Philpott 2017). It is also possible to combine a GLM with other methods, in this case a GAM, using ensemble modelling to improve results. This approach predicted possible changes in the future of biodiversity caused by climate change of grasshoppers, butterflies and dragonflies. Additionally, Maes et al. (2010) applied UK to interpolate climate data and support the main model's predictions. While GLMs are more commonly used in the literature to analyse and predict biodiversity, geostatistics and IDW are also applied for this purpose. Marchand et al. (2015) used IDW to interpolate the inflorescence density to support the analysis of bee foraging behaviour. Hortal and Lobo (2011) compared GLMs, using 24 predictor variables, and multivariate kriging to predict scarab dung beetle richness. Both models performed well on the training dataset but showed poor generalisation when applied to other locations, despite only minor differences in environmental conditions. One reason for the poor predictions could be the non-stationary nature of the variations in species abundance and the correlation to environmental variables. With better success, Ballesteros-Mejia et al. (2013) used CoK to map biodiversity patterns of

TABLE 6 | Summary of studies on biodiversity prediction using different methods.

Study	Methods	Species	Sampling time frame	Secondary variables[a]
Ballesteros-Mejia et al. (2013)	CoK	African sphingid moths (*Macroglossum trochilus*)	Not specified	Vegetation type, population density, infrastructure, protected areas
Hawkins and DeVries (2009)	GLM	Different butterfly species	Not specified	Evapotranspiration, temperature, forest biomes, vegetation index
Hortal and Lobo (2011)	GLM, kriging	Scarab dung beetles (Coleoptera Scarabaeidae)	Not specified	24 predictors accounting for soil, relief, climate, land use, habitat diversity and geographical location
Maes et al. (2010)	GLM, GAM, UK	Grasshoppers, butterflies, dragonflies	15 years	Temperature, rainfall, land cover, soil data
Marchand et al. (2015)	IDW	Bees	4 months	—
Monfared et al. (2013)	CoK, OK	Bees (Apoidea Hymenoptera)	19 months	Elevation, temperature
Plascencia and Philpott (2017)	GLM	Bees	4 days	Urban land cover, number of flowers, nearest neighbour analysis

[a] If applicable for the respective methods.

hawkmoths, using vegetation type, population density, level of infrastructure and protected areas as predictors. CoK yielded good results, but the results were biased towards densely populated areas due to a higher sampling density. This indicates that consistent sampling across regions is critical for such large-scale interpolation.

3 | Methodological Guidelines and Available Data

The application of spatial interpolation methods in entomology requires careful methodological considerations, including the selection of appropriate models, data collection strategies and validation techniques. In this section, we provide practical guidelines to assist researchers in choosing and implementing interpolation methods effectively. These guidelines are derived from the reviewed literature and are intended to help avoid common pitfalls in spatial data analysis.

In addition, we present a selection of publicly available datasets relevant to pest management, biodiversity analysis and disease vector mapping. These datasets serve as resources for researchers looking to test and validate interpolation techniques in real-world scenarios. The availability of well-documented data is essential for improving reproducibility and advancing methodological approaches in this field. By linking theoretical insights from the literature with concrete datasets, we aim to facilitate the application of spatial interpolation methods and support further research in insect ecology.

3.1 | Methodological Guidelines

Based on the reviewed literature, OK is the SIM of choice for local applications in pest management due to its robustness and sampling flexibility. This is especially relevant when remotely sensed data is unavailable at the local scale, as OK is the only interpolation method used across all applications (Sciarretta and Trematerra 2014). For effective semivariogram modelling, the literature suggests using 25–30 samples with approximately 4–5 lags (Journel and Huijbregts 1976). Although this recommendation originates from mining studies rather than entomology, it is frequently cited in research on spatial interpolation of insect-related data. However, some literature recommends having between 50 and 100 samples to improve spatial interpolation precision. Nevertheless, increasing the sample size generally leads to higher precision (Webster and Oliver 2007). When using OK, it is essential to consider that each area must be treated with equal importance during the sampling process, regardless of whether the area contains low- or high-density values (Brenner et al. 1997). The observed area should be constant and heterogeneous, ensuring that the spatial model can effectively capture the underlying variation (Zimmerman et al. 1999). OK does not require regular sampling, which is useful in the context of pest management, since sampling is often affected by the distribution of the crops (Schotzko and O'Keeffe 1989). Cross validation with several semivariogram models is important. Different semivariogram models can yield varying results; therefore, testing alternative models enhances the understanding of the spatial structure of the data and can improve the accuracy of predictions. OK is unsuitable for data exhibiting high variance across the study area; kriging methods able to model trends like UK are a better fit in such cases (Cressie 1986). OK can be implemented through several software platforms, with SURFER and ArcGIS being commonly used for geostatistical analysis. SURFER is a Geographic Information System (GIS) software package developed by Golden Software LLC, which supports various interpolation methods (e.g., kriging and IDW) and enables the creation of contour maps, heatmaps and 3D surface models. ArcGIS is a comprehensive GIS platform developed by Esri (Environmental Systems Research Institute Inc.) that supports spatial analysis, map creation and geostatistical tools such as kriging and IDW interpolation.

Different variations of CoK proof to be useful interpolation methods for pest management on a regional or global scale and are also employed for mapping vectors and biodiversity in several studies. Many of the same guidelines discussed for OK also apply to CoK. This is especially true regarding sampling strategies and spatial characteristics of the study area. Additionally, CoK is particularly useful when primary data is sparsely sampled, but auxiliary data is more readily available (Myers 1984). Correlation between the primary and secondary variables is critical for the spatial interpolation methods that incorporate auxiliary information (Goovaerts 1997). The effectiveness of CoK relies on selecting covariates that are strongly correlated with the variable of interest. The closer the relationship, the more accurate the predictions. This can significantly enhance prediction accuracy, often surpassing traditional kriging approaches (Stein and Corsten 1991). Auxiliary variables need to be sampled densely enough to model spatial correlations effectively. If auxiliary data is sparse or poorly sampled, the benefits of CoK diminish considerably. The same principle applies as with univariate kriging: secondary variables must be sampled in a manner that allows for spatial correlation to emerge (Knotters et al. 1995). Most studies utilise ArcGIS for CoK; in some cases, software packages in R are also used. R is a widely used programming language, developed by R Core Team, which provides a wide range of statistical and spatial analysis packages suitable for reproducible research.

In the context of disease vector mapping, no single preferential model emerges from the literature; however, SDMs are frequently applied. This suggests that multivariate and non-geostatistical models based on auxiliary data are useful for modelling the distribution of vectors. Sampling bias can significantly distort SDM predictions. A common source of bias stems from preferential sampling, where regions expected to host the species are disproportionately surveyed. To mitigate these effects, methods such as background sampling or spatial corrections should be employed to create a balanced representation of the landscape (Pennino et al. 2018). The number of samples required for accurate predictions in SDMs largely depends on the species' ecological niche and prevalence. Species with narrow distributions may require as few as 14 samples, whereas widespread species typically need 25 or more (Proosdij et al. 2016). It is essential that SDMs are tailored to the specific species of interest and the intended use of the model. Iterative development, testing, and validation help ensure the reliability of the predictions (Sofaer et al. 2019). When selecting an SDM, informed decisions should be made regarding model complexity, choice of predictor variables and the representation of predictions (e.g., via thresholds). More complex models do not necessarily yield better results. Simpler models, when chosen correctly, may offer superior performance by minimising overfitting and making predictions more robust (Merow et al. 2014). Most SDMs in the reviewed literature are implemented in R, using various R-packages. For certain models, such as MaxEnt, dedicated software applications are available, including the MaxEnt package itself.

GLMs are the most frequently used method for modelling biodiversity, as it is a complex influenced by numerous environmental variables. The process of variable selection is critical in building efficient GLMs (Hawkins and DeVries 2009). It is important to acknowledge that model performance varies; a particular model may perform poorly in some cases but yield robust predictions in others (Von Tress 2003). As Jovanovic (1991) highlights, careful selection of regression parameters and the use of diagnostic measures are essential to ensure the validity of model assumptions. Inefficient or improper variable selection can lead to poor model performance or biased predictions. The quality of spatial interpolation heavily depends on the design of the sampling plan. In the context of GLMs, spatial sampling should be designed to capture the full range of variation in covariates that affect the spatial process. Breslow and Clayton (1993). Selecting the appropriate GLM and its structure is crucial. Cross-validation and information-theoretic criteria, such as the Akaike information criterion (AIC), which balances model fit and complexity, should be used to guide model selection. Additionally, GLMs should be validated using independent datasets to confirm that they generalise effectively to unsampled areas (Varin et al. 2005). Besides being used as stand-alone models, GLMs can also be combined with kriging as regression kriging (RK) to enhance predictions, as they are themselves a form of regression (Liu et al. 2009). GLMs are mostly implemented using R and its statistical packages.

3.2 | Available Data

In this section, we present a curated selection of datasets relevant to the applications discussed in this review. These datasets were chosen because they contain representative, long-term observational data, making them well-suited for applied research and methodological testing of spatial interpolation techniques. Their extended collection periods and comprehensive spatial coverage allow for robust experimentation, helping to assess the performance of different interpolation methods in real-world ecological scenarios.

Table 7 provides an overview of the datasets, including the primary species observed, key features, intended application, geographic coverage, number of recorded entries and data collection time frame. For reference, each dataset is assigned a number from [1] to [8]. Each dataset includes spatial coordinates linked to species occurrence records, either as unique observations or as count data. In Table 7, datasets with singular observations are labelled as 'Observations', whereas those with varying counts per entry are categorised as 'Counts'. As many publicly available datasets focus mainly on species occurrences without additional environmental variables, they are often not directly suitable for

TABLE 7 | Summary of datasets on insect occurrence data.

Main species	Features	Application	Area	Entries	Time frame
[1] Monarch butterfly	Counts	Study of monarch migration phenology	North America	138,715	1996–present
[2] Bombus	Counts, elevation, sex, caste, host plant	Standardised survey of bee pollinators	USA	10,211	2007–2010
[3] Bees and wasps	Counts, flower/substrate, habitat, trap information	Aggregation of occurrence data	USA	530,000	1990–2019
[4] Mosquito	Counts, habitat, breeding site, water presence	Studying biodiversity and monitoring changes	Belgium	2346	2007–2011
[5] Mosquito	Observations	Vector monitoring, warning system	Europe	37,828	2014–2024
[6] Sand Flies	Counts	Data analysis with GIS software	São Paulo, Brazil	1164	2000–2007
[7] Box tree moths	Observations, observation type	Pest monitoring	England	36,480	2005–2021
[8] True bugs	Counts	Biodiversity analysis	Flanders, Brussels (Belgium)	96,249	2008–2018

SDM. However, they can serve as a foundational resource for integrating external data, such as climate variables, habitat information, or land cover data, to improve predictive modelling.

All datasets presented in this paper are openly accessible and can be downloaded in CSV format. Dataset sources and access links are listed in the bibliography.

Dataset [1] (Sheehan and Wever-Grullon 2021) contains occurrence records of monarch butterflies in North America, collected by community scientists as part of a long-term science programme initiated in 1996. Observers do not follow a standardised sampling protocol, nor do they conduct observations at fixed times or intervals, which may affect data quality. The dataset also provides links to the photograph, through which sightings are submitted. Next to butterflies, the dataset also includes sightings of milkweed, monarch eggs, larva and roost sightings. Dataset [2] (Koch 2021) documents Bombus abundance across the United States and was developed as part of a large-scale standardised survey of insect pollinators. The reason for the creation of this dataset was to document the decline and conservation status of Bombus species. It is the only dataset in this review that includes additional variables such as elevation, insect sex caste and host plant. Dataset [3] (Droege and Maffei 2023) aggregates data from numerous projects focused on bees and wasps across North America. Next to bees and wasps, it also includes other species like moths, butterflies, ants and more which are often collected during bee and wasp sampling. In some cases, observations include notes on the flower or substrate from which the species were collected. To validate the observations, each entry has been reviewed by project managers. Most public datasets on disease vectors focus on various mosquito species, including dataset [4] (Versteirt et al. 2022). The dataset stems from the Belgian project Modirisk by the Institute of Tropical Medicine studying and monitoring mosquitoes in Belgium due to the re-emergence of vector-borne diseases in Europe. With the dataset, the project aimed to predict changes in the biodiversity of mosquitoes to analyse the risk of diseases. The samples were taken using bait traps on corine land cover, urban nature and agriculture areas. All samples were double-checked to ensure data quality control. Dataset [5] (Mosquito Alert 2022) consists of data collected through the citizen science system Mosquito Alert. Each record includes a photograph and, if it concerns one of the five target mosquito species, is validated by four independent entomological experts. The project aims to serve as an early warning system for the detection of species potentially carrying vector-borne diseases. Dataset [6] (Shimabukuro et al. 2024) contains observations of sand flies in the São Paulo region of Brazil. The dataset compiles existing data on sand flies to support understanding of the distribution of American cutaneous leishmaniasis in the region. It was created to be used for data analysis using GIS software. It has a total of 1164 entries, 400 of these entries include individual count data. The entries are recorded from over 200 unique locations, and data sources include entomological collections in museums and health departments, and scientific publications. The first dataset [7] (Royal Horticultural Society 2023) on pest management consists of observations of the box tree moth in the UK by the Royal Horticultural Society. The dataset contains 36,480 entries, with the highest number recorded in 2021 (7268 entries). All records have been verified by entomologists. Another dataset [8] (Lock et al. 2021) covers true bugs (Hemiptera) occurrences in Belgium, specifically Flanders and the Brussels Capital Region. These northern Belgian regions are characterised by a mix of agricultural land and urban areas. 590 different species were submitted by volunteers, mainly since 2008. In 2018, the highest number of observations was submitted, totalling 21,455 entries. Next to the coordinates, the state of the entry is noted, which allows for tests on smaller scales, although the data is aggregated to 5×5 km grid cells. An additional source of insect data is the iNaturalist database, which compiles volunteer-submitted observations. For example, it includes 14,505 observations of the cotton bollworm, 12,943 observations of the Colorado potato beetle and 57,960 observations of the gypsy moth—species also covered in Section 2.3. The data is collected globally, and each entry includes geographic coordinates. With a free account, users can download the data as a CSV file for further analysis (iNaturalist 2024).

4 | Discussion

This study offers a comprehensive review and practical guidance on the application of spatial interpolation methods in insect ecology, with a particular emphasis on their relevance for ecological monitoring and pest management. The findings show that a wide variety of models are used across applications, including both univariate and multivariate kriging approaches, as well as various SDMs. Geostatistical techniques such as kriging have proven effective in modelling species distributions (Sciarretta and Trematerra 2014), though their accuracy strongly depends on data quality and sampling density (Webster and Oliver 2007). While useful, non-geostatistical methods often face challenges in capturing large-scale ecological complexities, highlighting a key area for improvement (Cressie 2015). A key takeaway of this review is the importance of comparing models before selecting an interpolation method. Even when using the same kriging method, such as OK, the choice of semivariogram models can determine whether spatial correlation is detected in the data (Park and Obrycki 2004). While these methods offer considerable practical utility, several challenges remain. The need for high-quality data and extensive sampling presents substantial costs and a key barrier to broader application in real-world contexts (Wen et al. 2009). This limitation could be partially addressed as sustainable agriculture and organic farming become more prominent (Sciarretta and Trematerra 2014). Moreover, current models often struggle to capture complex non-linear environmental interactions, indicating a need for further refinement to enhance predictive accuracy across diverse landscapes (Ballesteros-Mejia et al. 2013).

Emerging advances in artificial intelligence, machine learning (Du et al. 2020) and remote sensing (Comber and Zeng 2019) are poised to significantly influence the field of spatial interpolation. Although not yet widely established in entomology, these technologies offer promising avenues for improving spatial interpolation. For instance, various hybrid approaches combining machine learning and kriging have been proposed, including RK, kriging convolutional networks and neural network residual kriging (Zhu and Lin 2010; Appleby et al. 2020; Seo et al. 2015). These methods can improve predictions when spatial structures are not evident from point-based observations or when strong correlations exist between the target and secondary variables (Zhu and Lin 2010). Additionally, they handle large datasets more efficiently, accommodate diverse data distributions compared to conventional kriging, and perform well in contexts where spatial relationships are influenced by complex non-linear factors (Seo et al. 2015). However, these hybrid models often require greater computational resources and involve more parameters, which can make them more sensitive and challenging to optimise (Appleby et al. 2020). Furthermore, entirely machine learning-based spatial interpolation methods have recently emerged. These include conditional generative adversarial neural networks (Zhu et al. 2020), random forest spatial interpolation (Sekulić et al. 2020), a RFs variant that mimics kriging by incorporating spatial observations and distances from neighbouring points as covariates, and support vector machines (Du et al. 2020). These models can be transferred to entirely new datasets after training—an advantage not typically offered by kriging (Zhu et al. 2020). They also process data more efficiently and offer greater scalability (Sekulić et al. 2020). On the other hand, these machine learning approaches require a lot of training data, and the training process is complex, given that it also includes numerous parameters that must be defined. Moreover, non-uniform or random sampling strategies can reduce model performance (Du et al. 2020; Zhu et al. 2020). These technologies have the potential to reduce the need for extensive sampling while simultaneously enabling more accurate and dynamic ecological modelling (Chilès and Desassis 2018). Furthermore, incorporating spatio-temporal models could enhance the accuracy of predictions regarding insect population dynamics—crucial for addressing biodiversity loss and pest control in the face of climate change (Makori et al. 2017). In summary, although current spatial interpolation tools in insect ecology are robust, future advancements will likely favour more adaptive, data-driven models. Such advancements hold promise for significantly improving ecological monitoring and enabling more effective conservation and pest management strategies (Hengl 2007).

Author Contributions

Janne Heusler: conceptualisation, data curation, formal analysis; writing. **Jonas Funk:** conceptualisation, validation, review and editing. **Andreas Wagner:** conceptualisation, funding acquisition, project administration, resources, supervision, review and editing.

Acknowledgements

We would like to thank the Apicultural State Institute of the University Hohenheim for providing their specialised knowledge and their research premises. This paper and the BeeVision project are funded by the Carl-Zeiss-Stiftung as part of their 'CZS Wildcard' programme (https://www.carl-zeiss-stiftung.de/en/topics-projects/project-overview/detail/beevision-monitoring-der-bestaeuber-vielfalt-durch-dynamic-vision-sensorik). We acknowledge the use of AI tools to organise and articulate ideas, prepare code and find research sources during the preparation of this manuscript. Although the intellectual contributions and conceptual developments are entirely those of the authors, the role of AI tools in streamlining the writing process is clearly recognised. Open Access funding enabled and organized by Projekt DEAL.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

References

Agresti, A. 2019. *An Introduction to Categorical Data Analysis*. John Wiley & Sons.

Appleby, G., L. Liu, and L. P. Liu. 2020. “Kriging Convolutional Networks.” *Proceedings of the AAAI Conference on Artificial Intelligence* 34, no. 4: 3187–3194. https://doi.org/10.1609/aaai.v34i04.5716.

Azil, A. H., D. Bruce, and C. R. Williams. 2014. “Determining the Spatial Autocorrelation of Dengue Vector Populations: Influences of Mosquito Sampling Method, Covariables, and Vector Control.” *Journal of Vector Ecology* 39, no. 1: 153–163. https://doi.org/10.1111/j.1948-7134.2014.12082.x.

Ballesteros-Mejia, L., I. J. Kitching, W. Jetz, P. Nagel, and J. Beck. 2013. “Mapping the Biodiversity of Tropical Insects: Species Richness and Inventory Completeness of African Sphingid Moths.” *Global Ecology and Biogeography* 22, no. 5: 586–595. https://doi.org/10.1111/geb.12039.

Benndorf, J. 2023. *Angewandte Geodatenanalyse und -Modellierung: Eine Einführung in die Geostatistik für Geowissenschaftler und Geoingenieure*. Springer Fachmedien.

Bohling, G. 2005. “Introduction to Geostatistics and Variogram Analysis.” *Kansas Geological Survey* 1, no. 10: 1–20.

Brandhorst-Hubbard, J., K. Flanders, R. Mankin, et al. 2002. “Mapping of Soil Insect Infestations Sampled by Excavation and Acoustic Methods.” *Journal of Economic Entomology* 94: 1452–1458. https://doi.org/10.1603/0022-0493-94.6.1452.

Breiman, L. 2001. “Random Forests.” *Machine Learning* 45, no. 1: 5–32. https://doi.org/10.1023/A:1010933404324.

Breiman, L., J. Friedman, R. A. Olshen, et al. 2017. *Classification and Regression Trees*. Chapman and Hall/CRC.

Brenner, R., D. Focks, R. Arbogast, et al. 1997. “Practical Use of Spatial Analysis in Precision Targeting for Integrated Pest Management.” *American Entomologist* 44, no. 2: 79–102. https://doi.org/10.1093/ae/44.2.79.

Breslow, N., and D. Clayton. 1993. “Approximate Inference in Generalized Linear Mixed Models.” *Journal of the American Statistical Association* 88: 9–25. https://doi.org/10.1080/01621459.1993.10594284.

Bressan, A., F. J. Moral García, O. Sémétey, and E. Boudon-Padieu. 2010. “Spatio-Temporal Pattern of *Pentastiridius leporinus* Migration in an Ephemeral Cropping System.” *Agricultural and Forest Entomology* 12, no. 1: 59–68. https://doi.org/10.1111/j.1461-9563.2009.00450.x.

Callister, W. D., and D. G. Rethwisch. 2018. *Materials Science and Engineering: An Introduction*. Wiley.

Cano, J., M. Descalzo, N. Ndong-Mabale, et al. 2007. “Spatial and Temporal Variability of the *Glossina palpalis palpalis* Population in the Mbini Focus (Equatorial Guinea).” *International Journal of Health Geographics* 6: 36. https://doi.org/10.1186/1476-072X-6-36.

Castrignanò, A., L. Boccaccio, Y. Cohen, et al. 2012. “Spatio-Temporal Population Dynamics and Area-Wide Delineation of *Bactrocera oleae* Monitoring Zones Using Multi-Variate Geostatistics.” *Precision Agriculture* 13, no. 4: 421–441. https://doi.org/10.1007/s11119-012-9259-4.

Chilès, J. P., and N. Desassis. 2018. “Fifty Years of Kriging.” In *Handbook of Mathematical Geosciences: Fifty Years of IAMG*, edited by B. Daya Sagar, Q. Cheng, and F. Agterberg, 589–612. Springer International Publishing.

Cianci, D., N. Hartemink, and A. Ibánẽz-Justicia. 2015. “Modelling the Potential Spatial Distribution of Mosquito Species Using Three Different Techniques.” *International Journal of Health Geographics* 14: 10. https://doi.org/10.1186/s12942-015-0001-0.

Cocco, A., A. Q. Cossu, P. Erre, G. Nieddu, and P. Luciano. 2010. “Spatial Analysis of Gypsy Moth Populations in Sardinia Using Geostatistical and Climate Models.” *Agricultural and Forest Entomology* 12, no. 4: 417–426. https://doi.org/10.1111/j.1461-9563.2010.00488.x.

Comber, A., and W. Zeng. 2019. “Spatial Interpolation Using Areal Features: A Review of Methods and Opportunities Using New Forms of Data With Coded Illustrations.” *Geography Compass* 13, no. 10: e12465. https://doi.org/10.1111/gec3.12465.

Cressie, N. 1986. “Kriging Nonstationary Data.” *Journal of the American Statistical Association* 81: 625–634. https://doi.org/10.2307/2288990.

Cressie, N. 2015. *Statistics for Spatial Data*. John Wiley & Sons.

Cressie, N., and D. M. Hawkins. 1980. “Robust Estimation of the Variogram: I.” *Journal of the International Association for Mathematical Geology* 12, no. 2: 115–125. https://doi.org/10.1007/BF01035243.

Dale, M. R. T., and M. J. Fortin. 2014. *Spatial Analysis: A Guide for Ecologists*. Cambridge University Press.

De Cesare, L., D. E. Myers, and D. Posa. 2002. “FORTRAN Programs for Space-Time Modeling.” *Computers & Geosciences* 28, no. 2: 205–212. https://doi.org/10.1016/S0098-3004(01)00040-1.

Di Piazza, A., F. L. Conti, L. V. Noto, et al. 2011. “Comparative Analysis of Different Techniques for Spatial Interpolation of Rainfall Data to Create a Serially Complete Monthly Time Series of Precipitation for Sicily, Italy.” *International Journal of Applied Earth Observation and Geoinformation* 13, no. 3: 396–408. https://doi.org/10.1016/j.jag.2011.01.005.

Draper, N. R., and H. Smith. 1998. *Applied Regression Analysis*. John Wiley & Sons.

Droege, S., and C. Maffei. 2023. “Insect Species Occurrence Data From Multiple Projects Worldwide With Focus on Bees and Wasps in North America.” https://doi.org/10.15468/6AUTVB.

Du, P., X. Bai, K. Tan, et al. 2020. “Advances of Four Machine Learning Methods for Spatial Data Handling: A Review.” *Journal of Geovisualization and Spatial Analysis* 4, no. 1: 13. https://doi.org/10.1007/s41651-020-00048-5.

Duarte, F., M. V. Calvo, A. Borges, and I. B. Scatoni. 2015. “Geostatistics and Geographic Information Systems to Study the Spatial Distribution of *Grapholita molesta* (Busck) (Lepidoptera: Tortricidae) in Peach Fields.” *Neotropical Entomology* 44, no. 4: 319–327. https://doi.org/10.1007/s13744-015-0288-3.

Duncombe, J., A. Clements, J. Davis, W. Hu, P. Weinstein, and S. Ritchie. 2013. “Spatiotemporal Patterns of *Aedes aegypti* Populations in Cairns, Australia: Assessing Drivers of Dengue Transmission.” *Tropical Medicine & International Health* 18: 839–849. https://doi.org/10.1111/tmi.12115.

Eisen, L., and R. J. Eisen. 2011. “Using Geographic Information Systems and Decision Support Systems for the Prediction, Prevention, and Control of Vector-Borne Diseases.” *Annual Review of Entomology* 56: 41–61. https://doi.org/10.1146/annurev-ento-120709-144847.

Elith, J., S. J. Phillips, T. Hastie, M. Dudík, Y. E. Chee, and C. J. Yates. 2011. “A Statistical Explanation of MaxEnt for Ecologists.” *Diversity and Distributions* 17, no. 1: 43–57. https://doi.org/10.1111/j.1472-4642.2010.00725.x.

Ellsbury, M. M., W. D. Woodson, S. A. Clay, et al. 1998. “Geostatistical Characterization of the Spatial Distribution of Adult Corn Rootworm (Coleoptera: Chrysomelidae) Emergence.” *Environmental Entomology* 27, no. 4: 910–917. https://doi.org/10.1093/ee/27.4.910.

Espinosa, M., D. Weinberg, C. H. Rotela, F. Polop, M. Abril, and C. M. Scavuzzo. 2016. “Temporal Dynamics and Spatial Patterns of *Aedes aegypti* Breeding Sites, in the Context of a Dengue Control Program in Tartagal (Salta Province, Argentina).” *PLoS Neglected Tropical Diseases* 10, no. 5: e0004621. https://doi.org/10.1371/journal.pntd.0004621.

Espinosa, M. O., F. Polop, C. H. Rotela, et al. 2016. “Spatial Pattern Evolution of *Aedes aegypti* Breeding Sites in an Argentinean City Without a Dengue Vector Control Programme.” *Geospatial Health* 11, no. 3. https://doi.org/10.4081/gh.2016.471.

Farias, P. R. S., S. R. Roberto, J. R. S. Lopes, and D. Perecin. 2004. “Geostatistical Characterization of the Spatial Distribution of *Xylella* *fastidiosa* Sharpshooter Vectors on Citrus." *Neotropical Entomology* 33: 13–20. https://doi.org/10.1590/S1519-566X2004000100004.

Ferrier, S. 2002. "Mapping Spatial Pattern in Biodiversity for Regional Conservation Planning: Where to From Here?" *Systematic Biology* 51, no. 2: 331–363. https://doi.org/10.1080/10635150252899806.

Gallai, N., J. M. Salles, J. Settele, and B. E. Vaissière. 2009. "Economic Valuation of the Vulnerability of World Agriculture Confronted With Pollinator Decline." *Ecological Economics* 68, no. 3: 810–821. https://doi.org/10.1016/j.ecolecon.2008.06.014.

Giordano, B. V., K. W. Turner, and F. F. Hunter. 2018. "Geospatial Analysis and Seasonal Distribution of West Nile Virus Vectors (Diptera: Culicidae) in Southern Ontario, Canada." *International Journal of Environmental Research and Public Health* 15, no. 4: 614. https://doi.org/10.3390/ijerph15040614.

Goovaerts, P. 1997. *Geostatistics for Natural Resources Evaluation*. Oxford University Press.

Goulard, M., and M. Voltz. 1992. "Linear Coregionalization Model: Tools for Estimation and Choice of Cross-Variogram Matrix." *Mathematical Geology* 24, no. 3: 269–286. https://doi.org/10.1007/BF00893750.

Goulson, D., E. Nicholls, C. Botías, and E. L. Rotheray. 2015. "Bee Declines Driven by Combined Stress From Parasites, Pesticides, and Lack of Flowers." *Science* 347, no. 6229: 1255957. https://doi.org/10.1126/science.1255957.

Gribko, L. S., A. M. Liebhold, and M. E. Hohn. 1995. "Model to Predict Gypsy Moth (Lepidoptera: Lymantriidae) Defoliation Using Kriging and Logistic Regression." *Environmental Entomology* 24, no. 3: 529–537. https://doi.org/10.1093/ee/24.3.529.

Gringarten, E., and C. V. Deutsch. 2001. "Teacher's Aide Variogram Interpretation and Modeling." *Mathematical Geology* 33, no. 4: 507–534. https://doi.org/10.1023/A:1011093014141.

Hastie, T., and R. Tibshirani. 1990. "Exploring the Nature of Covariate Effects in the Proportional Hazards Model." *Biometrics* 46, no. 4: 1005–1016. https://doi.org/10.2307/2532444.

Hawkins, B. A., and P. J. DeVries. 2009. "Tropical Niche Conservatism and the Species Richness Gradient of North American Butterflies." *Journal of Biogeography* 36, no. 9: 1698–1711.

Hengl, T. 2007. *A Practical Guide to Geostatistical Mapping of Environmental Variables*. Institute for Environment and Sustainability.

Hohn, M., A. Liebhold, and L. Gribko. 1993. "Geostatistical Model for Forecasting Spatial Dynamics of Defoliation Caused by the Gypsy Moth (Lepidoptera: Lymantriidae)." *Environmental Entomology* 22, no. 1066: 1066–1075. https://doi.org/10.1093/ee/22.5.1066.

Hortal, J., and J. M. Lobo. 2011. "Can Species Richness Patterns Be Interpolated From a Limited Number of Well-Known Areas? Mapping Diversity Using GLM and Kriging." *Natureza & Conservação* 9. https://doi.org/10.4322/natcon.2011.026.

Hosmer, D. W., Jr., S. Lemeshow, and R. X. Sturdivant. 2013. *Applied Logistic Regression*. John Wiley & Sons.

iNaturalist. 2024. "Spatial Interpolation in Applied Insect Ecology: A Review, Including Guidelines and Datasets for Practical Use." https://www.inaturalist.org.

Isaaks, E. H., and R. M. Srivastava. 1989. *Applied Geostatistics*. Oxford University Press.

Janssen, S., G. Dumont, F. Fierens, and C. Mensink. 2008. "Spatial Interpolation of Air Pollution Measurements Using CORINE Land Cover Data." *Atmospheric Environment* 42, no. 20: 4884–4903. https://doi.org/10.1016/j.atmosenv.2008.02.043.

Journel, A. G., and C. J. Huijbregts. 1976. *Mining Geostatistics*. U.S. Department of Energy, Office of Scientific and Technical Information.

Jovanovic, B. D. 1991. "Linearization, Variable Selection and Diagnostics in Generalized Linear Models." PhD, University of Massachusetts, USA, uMI Order No. GAX91-20897.

Kemp, W., T. Kalaris, and W. Qulmby. 1989. "Rangeland Grasshopper (Orthoptera: Acrididae) Spatial Variability: Macroscale Population Assessment." *Journal of Economic Entomology* 82: 1270–1276. https://doi.org/10.1093/jee/82.5.1270.

Klink, R. v., T. August, Y. Bas, et al. 2022. "Emerging Technologies Revolutionise Insect Ecology and Monitoring." *Trends in Ecology & Evolution* 37, no. 10: 872–885. https://doi.org/10.1016/j.tree.2022.06.001.

Knotters, M., D. Brus, and J. Oude Voshaar. 1995. "A Comparison of Kriging, Co-Kriging and Kriging Combined With Regression for Spatial Interpolation of Horizon Depth With Censored Observations." *Geoderma* 67, no. 3–4: 227–246. https://doi.org/10.1016/0016-7061(95)00011-C.

Koch, F. H., and W. D. Smith. 2008. "Spatio-Temporal Analysis of *Xyleborus glabratus* (Coleoptera: Circulionidae: Scolytinae) Invasion in Eastern U.S. Forests." *Environmental Entomology* 37, no. 2: 442–452. https://doi.org/10.1603/0046-225X(2008)37[442:SAOXGC]2.0.CO;2.

Koch, J. B. 2021. "USBombus, Contemporary Survey Data of North American Bumble Bees (Hymenoptera, Apidae, Bombus) Distributed in the United States." https://doi.org/10.15468/G8CNKE.

Kravchenko, A., and D. G. Bullock. 1999. "A Comparative Study of Interpolation Methods for Mapping Soil Properties." *Agronomy Journal* 91, no. 3: 393–400. https://doi.org/10.2134/agronj1999.00021962009100030007x.

Li, J. 2008. *A Review of Spatial Interpolation Methods for Environmental Scientists*. Geoscience Australia.

Li, J., and A. D. Heap. 2014. "Spatial Interpolation Methods Applied in the Environmental Sciences: A Review." *Environmental Modelling & Software* 53: 173–189. https://doi.org/10.1016/j.envsoft.2013.12.008.

Liebhold, A., E. Luzader, R. Reardon, et al. 1998. "Forecasting Gypsy Moth (Lepidoptera: Lymantriidae) Defoliation With a Geographical Information System." *Journal of Economic Entomology* 91, no. 2: 464–472. https://doi.org/10.1093/jee/91.2.464.

Liebhold, A., R. Rossi, and W. Kemp. 2003. "Geostatistics and Geographic Information Systems in Applied Insect Ecology." *Annual Review of Entomology* 38: 303–327. https://doi.org/10.1146/annurev.en.38.010193.001511.

Liu, F., X. He, and L. Zhou. 2009. "Application of Generalized Regression Neural Network Residual Kriging for Terrain Surface Interpolation." In *International Symposium on Spatial Analysis, Spatial-Temporal Data Modeling, and Data Mining*, vol. 7492. https://doi.org/10.1117/12.837425.

Lock, K., F. Chérot, R. Libeer, et al. 2021. "Waarnemingen.be—Hemiptera Occurrences in Flanders and the Brussels Capital Region, Belgium." https://doi.org/10.15468/DKD0SG.

Maes, D., N. Titeux, J. Hortal, et al. 2010. "Predicted Insect Diversity Declines Under Climate Change in an Already Impoverished Region." *Journal of Insect Conservation* 14, no. 5: 485–498. https://doi.org/10.1007/s10841-010-9277-3.

Makori, D. M., A. T. Fombong, E. M. Abdel-Rahman, et al. 2017. "Predicting Spatial Distribution of Key Honeybee Pests in Kenya Using Remotely Sensed and Bioclimatic Variables: Key Honeybee Pests Distribution Models." *ISPRS International Journal of Geo-Information* 6, no. 3: 66. https://doi.org/10.3390/ijgi6030066.

Marchand, P., A. Harmon-Threatt, and I. Chapela. 2015. "Testing Models of Bee Foraging Behavior Through the Analysis of Pollen Loads and Floral Density Data." *Ecological Modelling* 313. https://doi.org/10.1016/j.ecolmodel.2015.06.019.

Martins, J. C., M. C. Picanço, R. S. Silva, et al. 2018. "Assessing the Spatial Distribution of *Tuta absoluta* (Lepidoptera: Gelechiidae) Eggs in Open-Field Tomato Cultivation Through Geostatistical Analysis." *Pest Management Science* 74, no. 1: 30–36. https://doi.org/10.1002/ps.4664.

McCullagh, P. 2019. *Generalized Linear Models*. Routledge.

Merow, C., M. J. Smith, T. Edwards, et al. 2014. "What Do We Gain From Simplicity Versus Complexity in Species Distribution Models." *Ecography* 37: 1267–1281. https://doi.org/10.1111/ECOG.00845.

Monfared, S., K. Khorshidian, A. Monfared, et al. 2013. "Prediction of Bees (Apoidea, Hymenoptera) Frequency and Distribution Using Best Linear Spatial Prediction Techniques (Kriging and Cokriging)." *Entomofauna* 34: 81–104.

Montgomery, D. C., and G. C. Runger. 2018. *Applied Statistics and Probability for Engineers*. Wiley.

Montgomery, G. A., M. W. Belitz, R. P. Guralnick, et al. 2021. "Standards and Best Practices for Monitoring and Benchmarking Insects." *Frontiers in Ecology and Evolution* 8. https://doi.org/10.3389/fevo.2020.579193.

Moral García, F. J. 2006. "Analysis of the Spatio–Temporal Distribution of *Helicoverpa armigera* Hb in a Tomato Field Using a Stochastic Approach." *Biosystems Engineering* 93: 253–259. https://doi.org/10.1016/j.biosystemseng.2005.12.011.

Mosquito Alert. 2022. "Mosquito Alert Dataset." https://doi.org/10.15470/T5A1OS.

Myers, D. 1984. "Co-Kriging—New Developments." In *Geostatistics for Natural Resources Characterization*, 295–305. https://doi.org/10.1007/978-94-009-3699-718.

Oliver, M. A., and R. Webster. 2015. *Basic Steps in Geostatistics: The Variogram and Kriging*. Springer International Publishing.

Park, Y. L., and J. Obrycki. 2004. "Spatio-Temporal Distribution of Corn Leaf Aphids (Homoptera: Aphididae) and Lady Beetles (Coleoptera: Coccinellidae) in Iowa Cornfields." *Biological Control* 31: 210–217. https://doi.org/10.1016/j.biocontrol.2004.06.008.

Pennino, M., I. Paradinas, J. Illian, et al. 2018. "Accounting for Preferential Sampling in Species Distribution Models." *Ecology and Evolution* 9: 653–663. https://doi.org/10.1002/ece3.4789.

Plascencia, M., and S. M. Philpott. 2017. "Floral Abundance, Richness, and Spatial Distribution Drive Urban Garden Bee Communities." *Bulletin of Entomological Research* 107, no. 5: 658–667. https://doi.org/10.1017/S0007485317000153.

Potts, S. G., J. C. Biesmeijer, C. Kremen, P. Neumann, O. Schweiger, and W. E. Kunin. 2010. "Global Pollinator Declines: Trends, Impacts and Drivers." *Trends in Ecology & Evolution* 25, no. 6: 345–353. https://doi.org/10.1016/j.tree.2010.01.007.

Potts, S. G., V. Imperatriz-Fonseca, H. T. Ngo, et al. 2016. "Safeguarding Pollinators and Their Values to Human Well-Being." *Nature* 540, no. 7632: 220–229. https://doi.org/10.1038/nature20588.

Proosdij, A. V., M. Sosef, J. Wieringa, et al. 2016. "Minimum Required Number of Specimen Records to Develop Accurate Species Distribution Models." *Ecography* 39: 542–552. https://doi.org/10.1111/ECOG.01509.

Reay-Jones, F. P. F. 2010. "Spatial and Temporal Patterns of Stink Bugs (Hemiptera: Pentatomidae) in Wheat." *Environmental Entomology* 39, no. 3: 944–955. https://doi.org/10.1603/EN09274.

Richards, S. L., C. S. Apperson, S. K. Ghosh, et al. 2006. "Spatial Analysis of *Aedes albopictus* (Diptera: Culicidae) Oviposition in Suburban Neighborhoods of a Piedmont Community in North Carolina." *Journal of Medical Entomology* 43, no. 5: 976–989. https://doi.org/10.1603/00222585(2006)43[976:saoaad]2.0.co;2.

Rijal, J. P., C. C. Brewster, and J. C. Bergh. 2014. "Spatial Distribution of Grape Root Borer (Lepidoptera: Sesiidae) Infestations in Virginia Vineyards and Implications for Sampling." *Environmental Entomology* 43: 716–728. https://doi.org/10.1603/EN13285.

Royal Horticultural Society. 2023. "RHS Box Tree Moth (*Cydalima perspectalis*) Monitoring." https://doi.org/10.15468/WG9L2V.

Ryan, P. A., S. A. Lyons, D. Alsemgeest, et al. 2004. "Spatial Statistical Analysis of Adult Mosquito (Diptera: Culicidae) Counts: An Example Using Light Trap Data, in Redland Shire, Southeastern Queensland, Australia." *Journal of Medical Entomology* 41, no. 6: 1143–1156. https://doi.org/10.1603/00222585-41.6.1143.

Schotzko, D. J., and L. E. O'Keeffe. 1989. "Geostatistical Description of the Spatial Distribution of *Lygus hesperus* (Heteroptera: Miridae) in Lentils." *Journal of Economic Entomology* 82, no. 5: 1277–1288. https://doi.org/10.1093/jee/82.5.1277.

Sciarretta, A., and P. Trematerra. 2014. "Geostatistical Tools for the Study of Insect Spatial Distribution: Practical Implications in the Integrated Management of Orchard and Vineyard Pests." *Plant Protection Science* 50: 97–110. https://doi.org/10.17221/40/2013-PPS.

Sciarretta, A., P. Trematerra, and J. Baumgärtner. 2001. "Geostatistical Analysis of *Cydia funebrana* (Lepidoptera: Tortricidae) Pheromone Trap Catches at Two Spatial Scales." *American Entomologist* 47, no. 3: 174–185. https://doi.org/10.1093/ae/47.3.174.

Sedda, L., L. Guerrini, J. Bouyer, N. Koné, and D. J. Rogers. 2010. "Spatio-Temporal Modelling of Glossina Palpalis Gambiensis and Glossina Tachinoides Apparent Densities in Fragmented Ecosystems of Burkina Faso." *Ecography* 33: 772–783. https://doi.org/10.1111/j.1600-0587.2009.06135.x.

Sekulić, A., M. Kilibarda, G. B. M. Heuvelink, M. Nikolić, and B. Bajat. 2020. "Random Forest Spatial Interpolation." *Remote Sensing* 12, no. 10: 1687. https://doi.org/10.3390/rs12101687.

Seo, Y., S. Kim, and V. P. Singh. 2015. "Estimating Spatial Precipitation Using Regression Kriging and Artificial Neural Network Residual Kriging (RKNNRK) Hybrid Approach." *Water Resources Management* 29, no. 7: 2189–2204. https://doi.org/10.1007/s11269-015-0935-9.

Sharov, A. A., E. A. Roberts, A. M. Liebhold, and F. W. Ravlin. 1995. "Gypsy Moth (Lepidoptera: Lymantriidae) Spread in the Central Appalachians: Three Methods for Species Boundary Estimation." *Environmental Entomology* 24, no. 6: 1529–1538. https://doi.org/10.1093/ee/24.6.1529.

Sheehan, N., and L. Wever-Grullon. 2021. "Journey North—Monarch Butterfly and Milkweed Observations by Volunteer Community Scientists Across Central and North America (1996–2020)." https://doi.org/10.6073/pasta/f7d7bef57f94b33b8a18a26954252412.

Shepard, D. 1968. "A Two-Dimensional Interpolation Function for Irregularly-Spaced Data." In *Proceedings of the 1968 23rd ACM National Conference*, 517–524. https://doi.org/10.1145/800186.810616.

Shimabukuro, P., T. R. da Silva, F. O. R. Fonseca, et al. 2024. "Fiocruz|Geographical Distribution of American Cutaneous Leishmaniasis and Its Phlebotomine Vectors (Diptera: Psychodidae) in the State of São Paulo, Brazil." https://doi.org/10.15468/SGWQ8P.

Sofaer, H., C. S. Jarnevich, I. Pearse, et al. 2019. "Development and Delivery of Species Distribution Models to Inform Decision-Making." *Bioscience* 69. https://doi.org/10.1093/BIOSCI/BIZ045.

Srivastava, V., A. D. Roe, M. A. Keena, R. C. Hamelin, and V. C. Griess. 2021. "Oh the Places They'll Go: Improving Species Distribution Modelling for Invasive Forest Pests in an Uncertain World." *Biological Invasions* 23, no. 1: 297–349. https://doi.org/10.1007/s10530-020-02372-9.

Stein, A., and L. Corsten. 1991. "Universal Kriging and Cokriging as a Regression Procedure." *Biometrics* 47: 575–587. https://doi.org/10.2307/2532147.

Vanbergen, A. J., and Insect Pollinators Initiative. 2013. "Threats to an Ecosystem Service: Pressures on Pollinators." *Frontiers in Ecology and the Environment* 11, no. 5: 251–259. https://doi.org/10.1890/120126.

Varin, C., G. Host, and O. Skare. 2005. "Pairwise Likelihood Inference in Spatial Generalized Linear Mixed Models." *Computational Statistics & Data Analysis* 49: 1173–1191. https://doi.org/10.1016/j.csda.2004.07.021.

Versteirt, V., W. Dekoninck, W. Van Bortel, et al. 2022. "MODIRISK: Monitoring of Mosquito Vectors of Disease (Inventory)." https://doi.org/10.15468/4FIDG2.

Vinatier, F., P. Tixier, P. F. Duyck, and F. Lescourret. 2011. "Factors and Mechanisms Explaining Spatial Heterogeneity: A Review of Methods for Insect Populations." *Methods in Ecology and Evolution* 2, no. 1: 11–22. https://doi.org/10.1111/j.2041-210X.2010.00059.x.

Voltz, M., and R. Webster. 1990. "A Comparison of Kriging, Cubic Splines and Classification for Predicting Soil Properties From Sample Information." *Journal of Soil Science* 41, no. 3: 473–490. https://doi.org/10.1111/j.1365-2389.1990.tb00080.x.

Von Tress, M. 2003. "Generalized, Linear, and Mixed Models." *Technometrics* 45, no. 1: 99. https://doi.org/10.1198/tech.2003.s13.

Wackernagel, H. 1995. *Multivariate Geostatistics*. Springer.

Webster, R., and M. A. Oliver. 2007. *Geostatistics for Environmental Scientists*. John Wiley & Sons.

Weisz, R., S. Fleischer, and Z. Smilowitz. 1995a. "Map Generation in High-Value Horticultural Integrated Pest Management: Appropriate Interpolation Methods for Site-Specific Pest Management of Colorado Potato Beetle (Coleoptera: Chrysomelidae)." *Journal of Economic Entomology* 88, no. 6: 1650–1657. https://doi.org/10.1093/jee/88.6.1650.

Weisz, R., S. Fleischer, and Z. Smilowitz. 1995b. "Site-Specific Integrated Pest Management for High Value Crops: Sample Units for Map Generation Using the Colorado Potato Beetle (Coleoptera: Chrysomelidae) as a Model System." *Journal of Economic Entomology* 88, no. 5: 1069–1080. https://doi.org/10.1093/jee/88.5.1069.

Wen, C., D. E. Guyer, and W. Li. 2009. "Local Feature-Based Identification and Classification for Orchard Insects." *Biosystems Engineering* 104, no. 3: 299–307. https://doi.org/10.1016/j.biosystemseng.2009.07.002.

Wong, D. W., L. Yuan, and S. A. Perlin. 2004. "Comparison of Spatial Interpolation Methods for the Estimation of Air Quality Data." *Journal of Exposure Science & Environmental Epidemiology* 14, no. 5: 404–415. https://doi.org/10.1038/sj.jea.7500338.

Zheng, X., D. Zhong, Y. He, and G. Zhou. 2019. "Seasonality Modeling of the Distribution of *Aedes albopictus* in China Based on Climatic and Environmental Suitability." *Infectious Diseases of Poverty* 8, no. 1: 98. https://doi.org/10.1186/s40249-019-0612-y.

Zhu, D., X. Cheng, F. Zhang, X. Yao, Y. Gao, and Y. Liu. 2020. "Spatial Interpolation Using Conditional Generative Adversarial Neural Networks." *International Journal of Geographical Information Science* 34, no. 4: 735–758. https://doi.org/10.1080/13658816.2019.1599122.

Zhu, Q., and H. S. Lin. 2010. "Comparing Ordinary Kriging and Regression Kriging for Soil Properties in Contrasting Landscapes." *Pedosphere* 20, no. 5: 594–606. https://doi.org/10.1016/S1002-0160(10)60049-5.

Zimmerman, D., C. Pavlik, A. Ruggles, et al. 1999. "An Experimental Comparison of Ordinary and Universal Kriging and Inverse Distance Weighting." *Mathematical Geology* 31, no. 4: 375–390. https://doi.org/10.1023/A:1007586507433.