BI4016 Research Project Thesis 2024/2025 School of Biological Sciences

Modelling to predict the population growth and cycles of *Microtus arvalis*

Student ID: 52100803

Work completed in partial fulfilment of the requirements for the degree BSc Honours Zoology

Word count: 5817

Acknowledgements

I would like to express my thanks to my project supervisor, Dr Deon Roos, for his guidance and support throughout. As well as making the project less intimidating, his insights and expertise have provided me with many valuable skills I will take into the future.

I would also like to thank Dr Constantino Caminero Saldaña for generously providing data that was crucial to the project. His willingness to assist me is greatly appreciated.

Abstract

Small rodent species commonly have complex population dynamics such as distinct outbreaks and cycles. These fluctuations can be highly disruptive, causing crop damage and spread of disease – issues which have been furthered by extreme management attempts and ongoing disagreement between farmers and conservationists regarding human-wildlife conflict. This study set out to develop a modelling approach to predict the population outbreaks of *Microtus arvalis* and identify which environmental characteristics could be used as predictors of population growth. A large dataset spanning nearly 30,000 km^2 of Northwest Spain, collected over almost 13 years was used for analysis ($n = 71{,}423$). A generalised additive model was used assuming a beta distribution with vole activity index (VAI) as the response variable, and 32 environmental characteristics as predictors. Model predictions produced a Root Mean Squared Error (RMSE) value of 0.13, which was higher than the mean of the response variable (0.07), showing the model was not predicting accurately. The RMSE was also calculated per pre-defined time period in order to measure how far in advance model predictions could be made. The model accurately predicted the VAI during time periods with no outbreak, but failed to predict the occurrence of an outbreak. Therefore, the model was found to be underpredicting actual values of VAI. Although the model was unable to predict outbreaks, the findings can still provide valuable insights into variables that contradict previous studies regarding predictive value of environmental variables such as temperature and soil type. The model's inaccuracy could be due to unusual outbreaks caused by climate change, or general ecological limitations. Suggestions were made concerning further work, specifically in terms of how the trialled methods could be improved.

Introduction

Cyclic population outbreaks are a well-documented occurrence in small rodent species across the globe. These cycles tend to be multiannual fluctuations in population size, typically progressing through four phases: increase, peak, crash, and low (Krebs and Myers, 1974). Some especially large peaks are classed as outbreaks, which vary substantially, making them more complex to study. The common vole (*Microtus arvalis*) serves as a notable example of a small rodent capable of both cyclic population dynamics and irregular outbreaks (Andreassen et al., 2020). During outbreaks, numbers of *Microtus arvalis* individuals can escalate dramatically, reaching up to 2,000 per hectare (Constantino Caminero-Saldaña et al., 2023).

Microtus arvalis becomes a significant pest when population numbers are high, having a negative impact on the surrounding environment. One of the primary ways *Microtus arvalis* affects agriculture is through crop damage. Rodents such as *Microtus arvalis* damage crops directly through grazing and burrowing in agricultural land, with grasslands, clover, alfalfa, winter cereals, and rape being the most at risk to vole damage in European agriculture (Singleton et al., 2010). Yield losses can be substantial, with rice crops down by 50% in some seasons in Asia (Singleton, 2003), and even up to 100% yield loss seen in Castilla y León (Deon Roos, Pers. Comm). Putting this into perspective, 5% of global rice production is enough to feed 180 million people for a year (Singleton, 2003), revealing the threat to food security. In 2007, a severe outbreak of *Microtus arvalis* in Central Spain caused the largest recorded losses in cereal crops, potatoes, and vineyards, resulting in management costs of approximately €15 million—an amount that the government was unable to fully compensate for farmers (JCYL 2008). More recently, a 2021 assessment using drone footage revealed that *Microtus arvalis* had damaged a quarter of the alfalfa fields surveyed during the year (Plaza et al., 2021). Beyond direct crop losses, secondary costs can also arise for farmers. This can include ploughing degraded soil, replanting crops, and repairing damaged infrastructure (Singleton et al., 2010). Due to the frequent and extensive damage they cause, *Microtus arvalis* are considered the most significant vertebrate pest in European agriculture (Lauret et al., 2019).

Beyond causing damage to crops, *Microtus arvalis* can also spread disease. This is attributed to the spread of the bacterium *Francisella tularensis* within large vole populations. As the bacterium spreads rapidly throughout these groups, it triggers outbreaks of tularemia, a serious infectious disease with substantial impacts on surrounding organisms (Herrero-Cófreces et al., 2021). A previous study found a positive correlation between *Francisella tularensis* prevalence and vole population density (Rodríguez-Pastor et al., 2017). During a population peak, 33% of sampled voles were infected, demonstrating their capacity to transmit tularemia at high rates during such periods (Rodríguez-Pastor et al., 2017). Notably, increased tularemia transmission to humans has also been linked to *Microtus arvalis* populations in northwestern Spain (Luque-Larena et al., 2015). Human outbreaks tend to coincide with the peak density phases of vole population cycles, posing public health risks in these areas (Herrero-Cófreces et al., 2021). The presence of *Microtus arvalis* and its role in spreading tularemia also negatively affects local ecosystems. For example, the Iberian hare (*Lepus granatensis*), an endemic species in northwestern Spain, has experienced population growth rate declines due to increased mortality caused by the bacterium *Francisella tularensis* during *Microtus arvalis* outbreaks (Rouco et al., 2023). Additionally, *Lepus granatensis* is vital to the survival of critically endangered predators in the area, including the Spanish imperial eagle (*Aquila adalberti*) and the Iberian lynx (*Lynx pardinus*). Consequently, fluctuations in *Microtus arvalis* populations could significantly harm biodiversity over time. As this issue is pronounced during outbreaks of *Microtus arvalis,* preventing population growth at the beginning of an outbreak is necessary to minimise the problem. However, halting growth at this stage would require insight into the timing and location of the outbreak – such as a method of prediction.

The risks posed by rodent population fluctuations have led to some extreme control measures. In 1998, a bounty scheme was introduced in Southeast Asia to combat rodent damage, particularly to rice crops. The program incentivised people to capture and kill rats using various methods, including bamboo snares, caging, and electrocution (Singleton et al., 2007). However, due to limited biological monitoring in many developing countries, numerous species are classified as "insufficiently known" in terms of conservation status (Aplin et al., 2006). The bounty scheme, which rewarded payments based on the collection of rat tails without consideration for species identification, posed risk to such wildlife (Singleton et al., 2007). In 2009, it was believed an outbreak of plateau pikas (*Ochotona curzoniae*) in China caused the spread of desert land. To control the outbreak, contraceptive pellets were spread across large areas of habitat to discourage breeding. However, researchers outside of China believed it was wrongly implemented, disputing the cause of the desertification, and conservationists argued it may unnecessarily harm non-target animals and predators. The disagreement suggests the use of contraceptive pellets in this context represents a measure which was too extreme for the understanding of the organisms at the time, which was implemented fast amid fears of the damage the pikas may have been causing (Watts, 2009). Chemical control is a common measure. In an outbreak in northwest Spain between 2006 and 2007, rodenticide was applied to crops extensively by government authorities. This was thought to take place without enough research into the effects on crops or other wildlife. Consequently, some of the crops were ruled toxic for human consumption, resulting in alleged reduced product and increased mortality rate for many non-target species (Olea et al., 2009).

Management strategies aimed at addressing these issues are also challenging due to the complexities of human-wildlife conflict. Sustainable management is made unattainable by clashing views surrounding causes and management of outbreaks between conservationists and farmers (Lauret et al., 2019). Intense disputes from conservationists arise when severe practices such as large-scale chemical use take place, putting non-target species at risk (Olea et al. 2009). On the other hand, farming associations tend to believe the conservationists decisions threaten their identity and their industry (Lauret et al., 2019). In a survey carried out in Castilla y León, conservationists suggested *Microtus arvalis* outbreaks to be caused by existing agricultural practices, whereas farmers generally disagreed. Additionally, governmental agencies thought the problem lay on the relationship between the two groups (Lauret et al., 2019). After a *Microtus arvalis* outbreak in Castilla y León in 2014, rodenticide had been prohibited, however, it is still considered necessary beyond the use of preventative measures (Lauret et al., 2019). This shows there are social tensions present that current strategies are yet to overcome, adding further complexity to the issue.

The challenges of implementing responsive management strategies emphasise the importance of prioritising prediction over reaction, allowing potential problems to be addressed proactively before they escalate. Additionally, for prediction to be useful in implementing management strategies, it should be made as far in advance as possible, ideally with at least a year's notice (Caminero-Saldaña, Pers. comm). In Castilla y León, methods for predicting when and where *Microtus arvalis* populations are going to fluctuate are yet to be developed. More broadly, based on available information, there are no current techniques for predicting fluctuations in species with complex population dynamics. Therefore, this study aimed to develop a modelling approach to predict the population growth and cycles of *Microtus arvalis,* and identify which environmental characteristics could be used as predictors of population growth. A previous study found that within the complex nature of the cyclic population dynamics of *Microtus* *arvalis* in Castilla y León, synchronising influences and dispersal of limiting factors operate across both space and time (Roos et al., 2022). Accordingly, the current study also aimed to produce a model which accounts for temporal and spatial trends. Shared mechanisms have been proposed for *Microtus arvalis* population cycles across various locations, indicating that new approaches may have the potential to be globally transferrable (Lambin et al., 2006), and additionally could be applicable to other species.

Methods

Data

This study used data collected and provided by the Instituto Tecnológico Agrario de Castilla y León between winter of 2011 and summer of 2024. The Castilla y León region is characterized by a flat plateau surrounded by mountain ranges, and split by the Duero River. Starting in the 1970s, human-driven agricultural land-use changes such as crop cultivation and irrigation facilitated the movement of *Microtus arvalis* populations from the mountain areas to the dryer lowlands (Herrero-Cófreces et al., 2021). As a result, vole densities were higher on the plateau compared to the surrounding highlands (Jareño et al., 2015). Following on, crop damage patterns indicate *Microtus arvalis* had outbreaks every five years since then (Luque-Larena et al., 2013). To track vole populations efficiently, practitioners used a standardized abundance index based on vole activity indicators. Surveys were carried out along transects following linear landscape features such as field edges, track margins, and ditch banks, all known to serve as habitats for voles. Each transect was systematically divided into smaller units, to document evidence of vole presence such as latrines, freshly cut vegetation, and new burrows. The abundance index was calculated as the proportion of these units showing signs of vole activity (Roos et al., 2022). The vole activity index (VAI), measured according to a continuous scale of 0–100%, was used as the response variable in this study (Figure 1).

Predictors

In addition to VAI, environmental variables that were considered likely to contribute to the occurrence of outbreaks were also gathered and were used as predictors for analysis (Figure 2). Previously, long-term weather trends have been found to affect populations of *Microtus arvalis,* including their range expansion (Jareño et al., 2015). Weather parameters were also found to have a strong connection to the outbreak risk of *Microtus arvalis,* and the thresholds of complex weather conditions correlate directly with population size (Esther et al., 2014). This has been seen especially during periods of very high or low population (Imholt et al., 2011). This relationship suggests weather may be a strong predictor of VAI (Esther et al., 2014). Climate variables such as precipitation and temperature are believed to have a positive relationship with rodents due to increasing food, and impacting reproduction (Li et al., 2020, Daketse and Martinet, 1977). These previous results lead precipitation and temperature to be included in the model (Figure 2A and B).

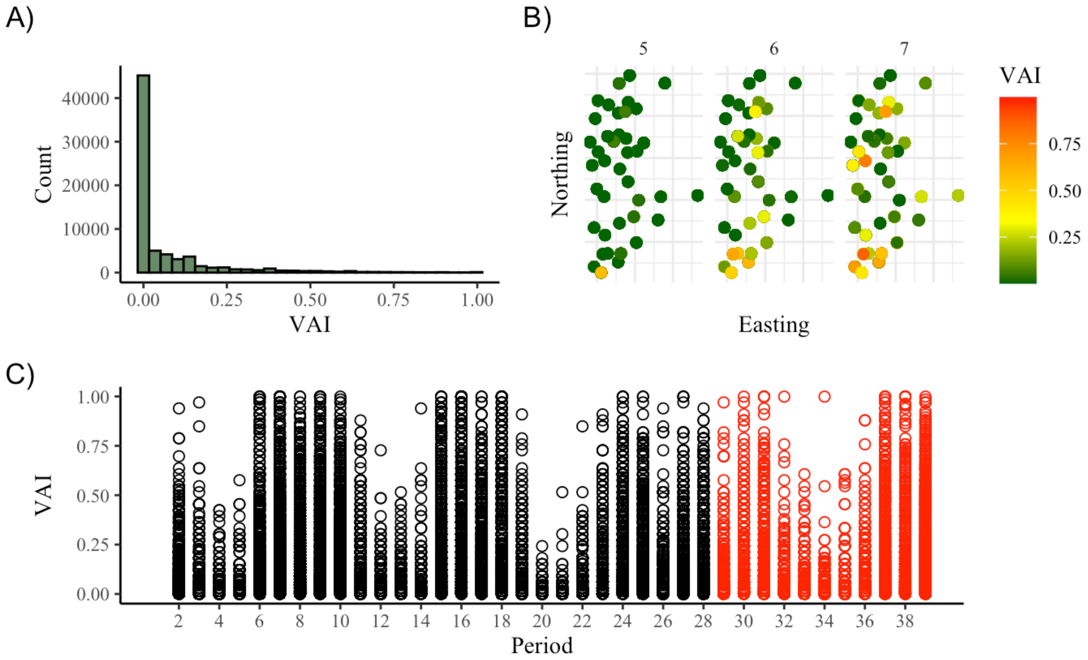

Figure 1. Vole activity Index (VAI) of *Microtus arvalis* in Castilla y León. A) The distribution of the frequency of each recorded value of VAI. B) Population groups plotted spatially across three exemplary four monthly time periods. Colour corresponds to the VAI. C) VAI plotted at all 39 four monthly time periods present in the data. Red points represent test data sectioned off in a train/test split.

Differences in vole population numbers have been found to coincide with different crop types including difference between annual or perennial (Heroldová et al., 2021), herbaceous or wooded plants (Jareño et al., 2015), and preference for specific crops, such as alfalfa (Santamaría et al., 2019). This highlighted crop information as a factor that could be used as a predictor, so nine different crop type variables were included in the modelling. More so than crop influences, agricultural practices are thought to affect population growth due to *Microtus arvalis* habitat use (Roos, 2019). Irrigation has strong associations with *Microtus arvalis* dynamics (Jareño et al., 2015), which was also included in the model (Figure 2C). Soil property and topography of the area have been found to be connected to *Microtus arvalis* outbreak (Blank et al., 2011). Therefore, altitude as well as different soil types and properties were included in the study.

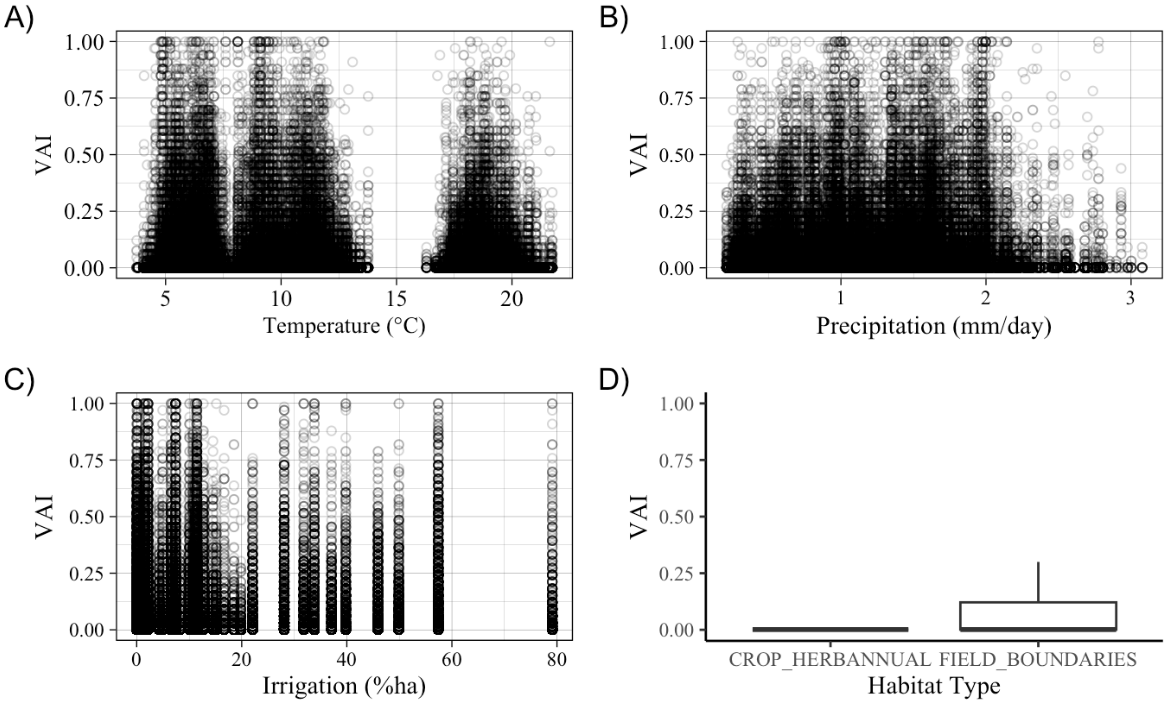

Figure 2. Relationships between vole activity index (VAI) of *Microtus arvalis* environmental variables. Showing VAI and A) temperature (°C) , B) precipitation (mm/day), C) irrigation (%ha) and D) habitat type (either herb annual crop or field boundaries).

Another relevant variable regards habitat, and whether vole activity is found in field margins or within crops, as previous research suggests *Microtus arvalis* abundance tends to be higher in field margins (Rodríguez-Pastor et al., 2016, Fotheringham et al., 2024) (Figure 2D). The connectivity of habitats and structural elements within them has also been explored, which influence *Microtus arvalis* outbreaks indirectly through affecting predators and species competition (Fotheringham et al., 2024). This led to data on dispersion pathways, and other geographical structures within areas such as roads or rivers being included in the model.

In summary, the predictor variables are all supported by biological justification and empirical evidence which warrants their inclusion. This suggests that, with a large sample size covering a large spatial and temporal scale, a predictive model should reveal the same fundamental relationships and accurately predict vole population size. Any deviation from this could be argued to raise questions about the usability, going forward, of previous research results in this area. All predictors are shown in the below equation.

Analysis

All data processing and analysis was carried out in R (R Core Team, 2023), with use of *mgvc* (Wood, 2011), *lubridate* (Grolemund and Wickham, 2011), *gratia* (Simpson, 2024) and *ggplot2* (Wickham, 2016) packages. The variables were processed and merged into a single dataset (see file “Supplementarymaterial.html”, uploaded separately). The final dataset consisted of 40 biologically relevant variables and 71,423 observations. An exploratory data analysis was carried out to understand the overall patterns of the data and check for invalid entries. As many of the predictors being used had the potential to have non-linear relationships with the VAI, smoothers were used, and were implemented using a generalised additive model (GAM) with the bam function and fREML (fast restricted maximum likelihood) in *mgvc* (Wood, 2011). A beta distribution, with corresponding logit link function, was assumed as VAI is a proportion bound between 0 and 1. The model accounted for space and time independently, using spatial coordinates (UTM) and temporal periods (four month increments). 32 variables were included in the model as explanatory variables. The number of basis functions (K) for each smooth term was adjusted to allow enough flexibility to fully represent the relationship, while still being biologically plausible. Thirty percent of the data was set aside for testing the model's predictive accuracy. Instead of the traditional random 70:30 split, the final 30% of the time series was used as the test data. This approach was chosen to avoid artificially high prediction accuracy that could arise from spatially and temporally correlated data. By splitting the data chronologically, the influence of the temporal and spatial autocorrelation was minimised. The model was therefore trained on the earlier 70% of the data and evaluated on the final 30%. Chronological splitting also allowed for the models accuracy to be evaluated over time, per sequential time period of the test data, ultimately allowing the test of its suitability to be used in realistic management strategies requiring advanced planning. See file “Supplementarymaterial.html” for detailed data analysis carried out.

The model corresponded to the following equation:

$$
\begin{aligned}
&y_{i,t} \sim Beta(\mu_{i,t}, \theta) \\
logit(\mu_{i,t}) = {} & \alpha_0 + f_1(Easting_i,) + f_2(temp_{i,t}) + f_2(maxtemp_{i,t}) + f_2(mintemp_{i,t}) \\
& + f_2(precip_{i,t}) + f_2(waterfc_i) + f_2(permeability_i) + f_2(altitude_i) \\
& + f_3(roadl_i) + f_3(waterl_i) + f_3(pathsl_i) + f_3(waterS_i) + f_3(roadS_i) \\
& + f_3(pathS_i) + f_3(waterw_i) + f_3(roadw_i) + f_3(pathw_i) \\
& + f_3(woodycrop_i) + f_3(annualcrop_i) + f_3(multiannualcrop_i) \\
& + f_3(treepasture_i) + f_3(bushypasture_i) + f_3(woodypasture_i) \\
& + f_3(herbpasture_i) + f_3(totalpasture_i) + f_3(fallow_i) + f_3(irrigation_i) \\
& + f_3(clay_i) + f_3(silt_i) + f_3(sand_i) + f_3(organic_i) + f_3(pH_i) \\
& + \alpha_1 \times fieldboundary_i
\end{aligned}
$$

where y is assumed to follow a $Beta$ distribution, with a mean μ and shape ϕ, and where μ is regressed on the $logit$ link as a function of location i in period t through a two-way non-linear interaction (f_1) between $Easting$ (X UTM) and $Northing$ (Y UTM) coordinates using a two dimensional thin plate regression spline with 10 knots in each dimension. In addition the model includes various f_2 functions, which used a thin plate regression spline with 10 knots to fit the relationships for mean temperature ($temp$, measured in degrees Celsius, per period), maximum temperature ($maxtemp$, measured in degrees Celsius, per period), minimum temperature ($mintemp$, measured in degrees Celsius, per period), precipitation ($precip$, measured as mean mm per day per period), water field capacity ($waterfc$, measured as % of soil water capacity per field), permeability ($permeability$, measured as mean mm/day in soil per area), and altitude ($altitude$, measured as mean meters above sea level per area). The model also includes f_3 functions, which used a thin plate regression spline with 5 knots to fit the relationships for the road length ($roadl$, measured as mean km/ha per area), water length ($waterl$, measured as mean km/ha per area), dispersion pathway length ($pathsl$, measured as mean km/ha per area), water surface ($waterS$, measured as mean %ha per area) road surface ($roadS$, measured as mean %ha per area), dispersion pathway surface ($pathS$, measured as mean %ha per area), water width ($waterw$, measured as mean meters per area), road width ($roadw$, measured as mean meters per area), dispersion path width ($pathw$, measured as mean meters per area), woody crop ($woodycrop$, measured as mean %ha per area), annual crop (a$nnualcrop$, measured as mean %ha per area), multiannual crop ($multiannualcrop$, measured as mean %ha per area), tree pasture ($treepasture$, measured as mean %ha per area), bushy pasture ($bushypasture$, measured as mean %ha per area), woody pasture ($woodypasture$, measured as mean %ha per area), herb pasture ($herbpasture$, measured as mean %ha per area), fallow ($fallow$, measured as mean % per ha) irrigation ($irrigation$, measured as mean % per ha), clay ($clay$, measured as mean % of soil per area) silt ($silt$, measured as mean % of soil per area), sand ($sand$, measured as mean% of soil per area), organic matter ($organic$, measured as mean % of soil per area) and pH (pH, measured as mean pH of soil per area). $\alpha 1$ was the contrast treatment for the binary field boundary ($fieldboundary$) compared to the intercept $\alpha 0$ (representing herb annual crop).

Results

Following the model being run on the training data complete with the aforementioned parameters, individual predictor relationships were observed using the gratia package (Simpson, 2024). Examples can be seen in Figure 3. Following the initial observations, the predictions produced by the model were compared with the actual values of the training data (Figure 4).

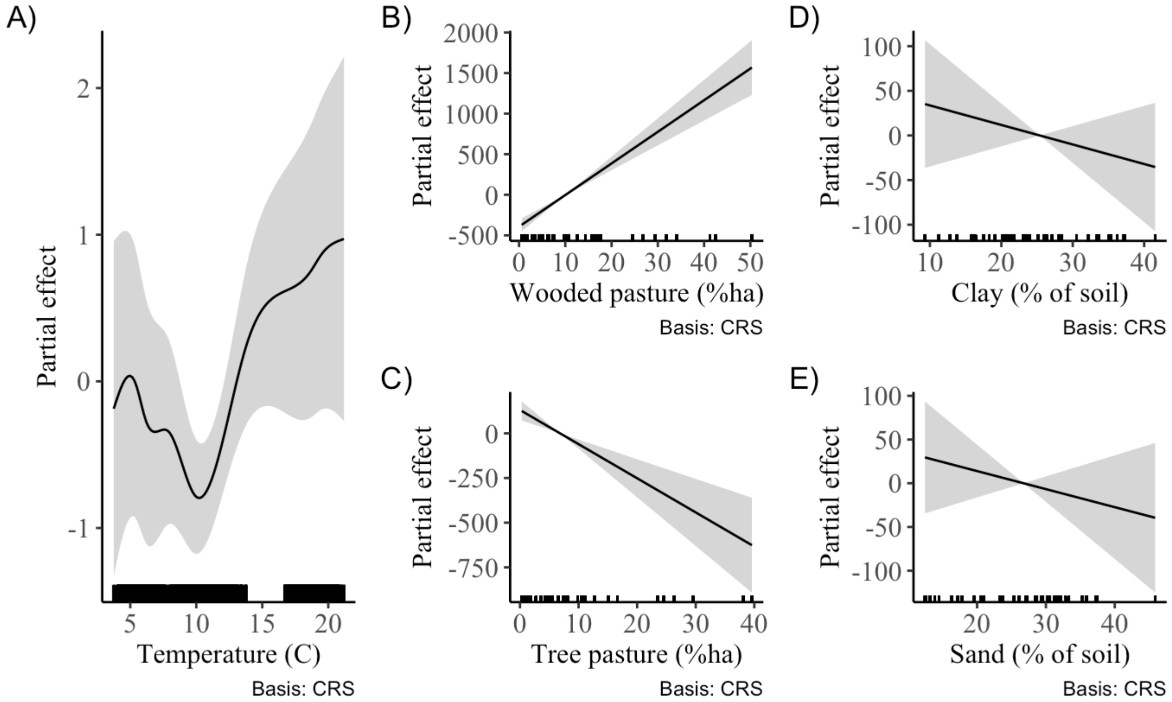

Figure 3. Relationships predicted between vole activity index (VAI) and environmental variables. Partial effect refers to the effect of the predictor on the response variable (VAI) when all other predictors are held constant. A) VAI and temperature. The prediction shows a largely negative relationship until just over 10 degrees, followed by a positive relationship until around 15 degrees, then levelling off slightly. B) VAI and wooded pasture (%ha). The prediction shows a largely positive relationship. C) VAI and tree pasture (ha%). The prediction shows a largely negative relationship. D) VAI and clay (% of soil). The prediction shows a slight negative relationship. E) VAI and sand (% of soil). The prediction shows a slight negative relationship.

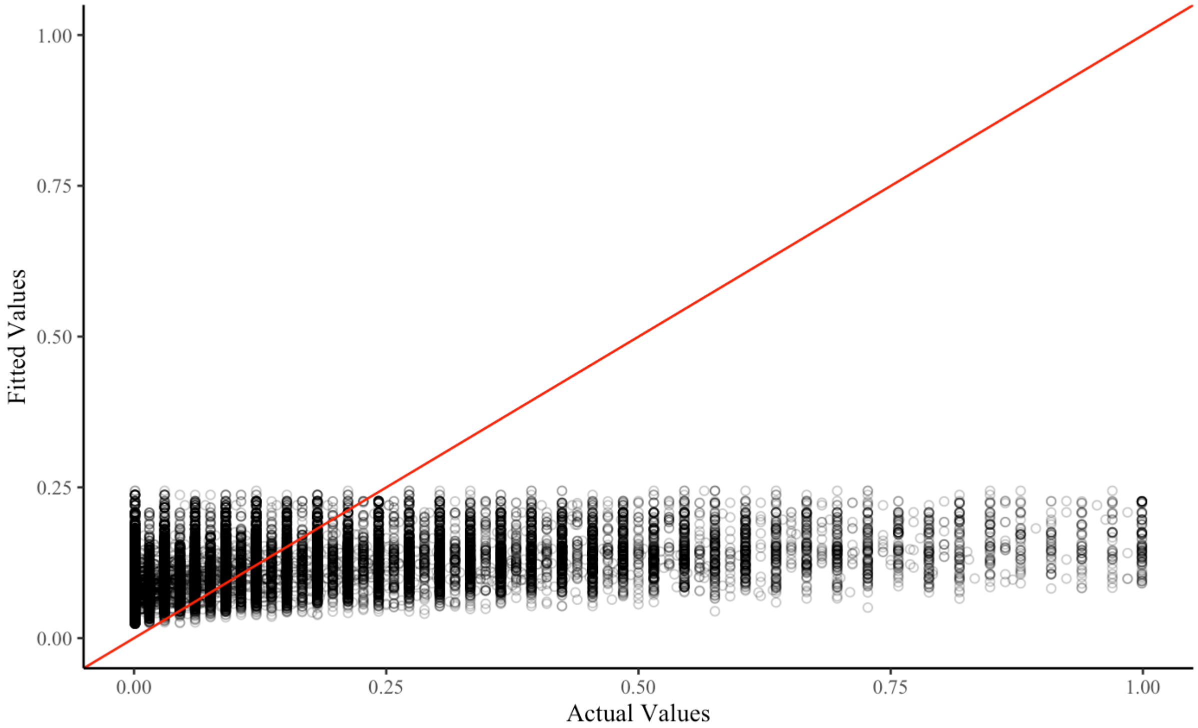

Figure 4. Predictions on training data. The relationship between predictions of *Microtus arvalis* population growth produced by a model (Fitted Values) and actual values from the training data. The actual values have a maximum of 1, while the fitted values have a maximum of 0.24.

The model was found to explain 21% of the deviance in VAI, however, the model underpredicts actual values of VAI, as indicated in Figure 4. Following the observation of training data predictions in Figure 4, the model was used to make predictions of VAI on the test data (i.e., out-of-sample VAI) – equating to periods 29-39 (Figure 5).

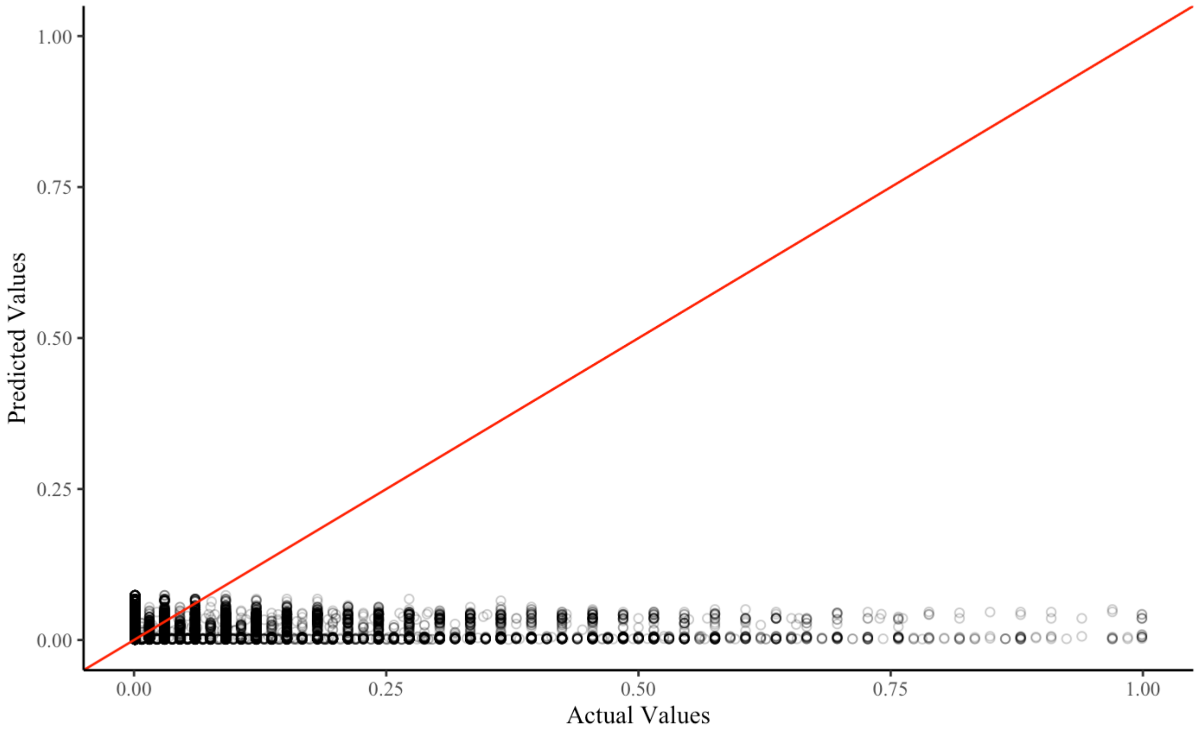

Figure 5. Predictions on test data. The relationship between predictions of *Microtus arvalis* population growth produced by a model and actual values from comparable test data. Predictions all lie within the range of 0 to 0.07, whereas actual values range from 0 to 1.

Model accuracy was evaluated according to Root Mean Squared Error (RMSE). Testing model predictions on the test data, the RMSE value was 0.13, which is higher than the mean of the response variable (0.07). Further, RMSE values closer to 0 indicate greater model fit. To test how far in advance predictions could be made, the test data was split into its pre-defined time periods and the RMSE was re-calculated for each (Figure 6).

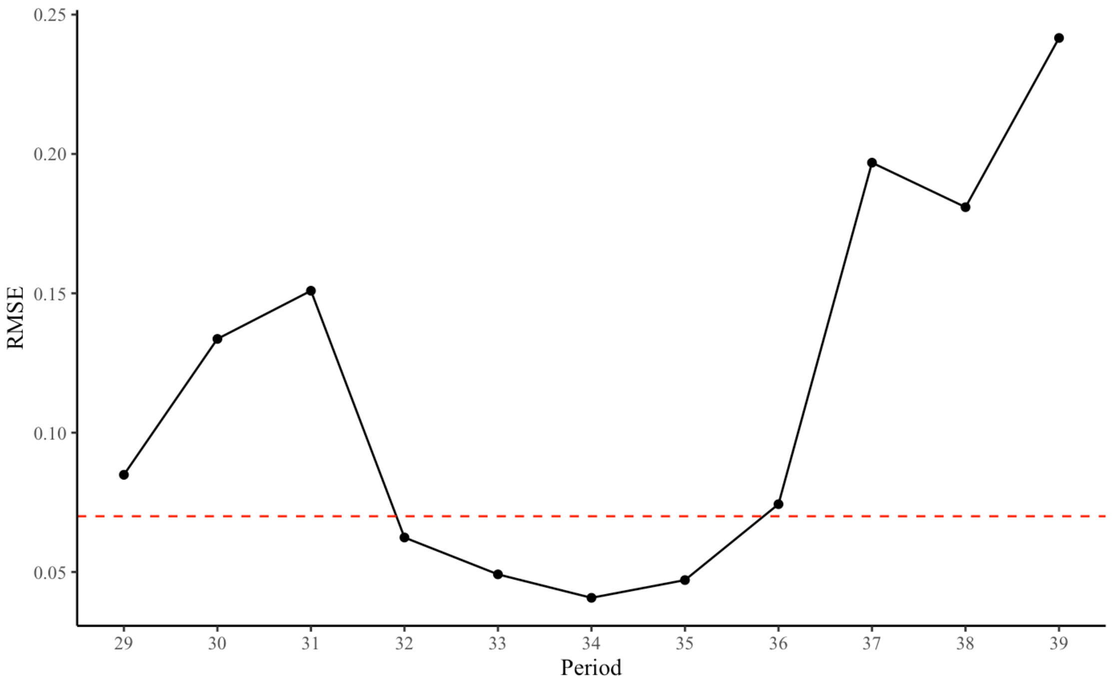

Figure 6. Predictive accuracy plotted temporally. The predictive accuracy of a *Microtus arvalis* population growth model, in different time periods. Calculated using Root Mean Squared Error (RMSE). The red line represents the mean vole activity index (0.07). The model is more accurate in periods 32 to 35 (as RMSE < 0.07), and less accurate before and after.

Discussion

In this paper, a predictive modelling approach was developed using a variety of environmental characteristics to predict the population growth of *Microtus arvalis*: such an approach can provide insight into when and where an outbreak will take place. Predictions were observed from the resulting model on the train and test data, as well as some individual predictor relationships. The model accounted for 21% of the variation in VAI, and had an accuracy of 0.13 (according to RMSE). Additionally, model predictions were found to be more accurate in some time periods (32-35) than others.

In Figure 4, the model's predictions on the training data are displayed. The points plotted diverge from the red line (Figure 4), suggesting that the predictions of VAI differ from the actual values. On the other hand, Figure 4 also shows arguably positive aspects: the model avoids predicting high values at the lowest actual values and predicts many zeroes, which aligns with the original data distribution (Figure 1A). To put this into context with the planned use of the model – predicting future outbreaks – predictions were also evaluated using the test data (Figure 5) and compared with the corresponding results from the training data (Figure 4). As expected, the model performance was superior using the training data (in-sample predictions) compared to the test data (out-of-sample predictions). Upon first inspection of Figure 5, it is clear that the model is underpredicting, with predicted values consistently low. While the model made predictions up to 0.24 on the training data, the highest predictions on the test data were only 0.07. Although this is considerably lower, the model's underperformance in Figure 5 must be weighed up against its overall potential, as shown in Figure 4, where the model's effectiveness was still constrained.

The deviance explained of the model was 21%, indicating that the model failed to explain most of the variance in the VAI. This is not necessarily a significant defect of the modelling approach, as exogenous influences may negatively affect predictive accuracy and population dynamics can often be too complex to explain fully. A notable reason for this is due to the inability to control ecological systems. Field environments show stark contrast to ideal settings for empirical research such as a controlled laboratory. They contain a wide range of diverse behaviours, interactions, and dynamics, many of which are not easily accounted for, often leading to a gap between theoretical predictions and practical outcomes (Beckage et al., 2011).

This complexity of prediction in realistic ecological systems suggests 21% deviance explained may be better than implied at first glance.

According to RMSE, the model's accuracy was calculated to be 0.13. This was compared against the mean VAI as a measure of being close enough to 0 for the model predictions to be deemed accurate. The RMSE (0.13) was only 0.06 higher than the VAI mean (0.07), so the model could have been seen to be close to performing successfully. However, Figure 5 showed the model was underpredicting on the test data. For example, for all points where the actual value was 1, meaning 100% of the sampled area showed signs of vole activity, the predicted values were no more than 0.05. Calculating the RMSE for each period (Figure 6) shows how far in advance an outbreak could be predicted. Ideally for management purposes, this would be one year in advance (Caminero-Saldaña, Pers. comm). Figure 6 shows that the model predictions were more accurate for periods 32 to 35 as RMSE < 0.07 (mean VAI), while the model performance was also competent during periods 29 and 36 (RMSE close to 0.07), while the other periods were not predicted accurately. This equates to the model being able to predict *Microtus arvalis* population numbers 16–28 months in advance, but not before or after. However, when considering actual values of VAI in the test data (Figure 1C), the periods in which the model is predicting accurately are also the periods of lowest VAI. This ties back to Figure 5 where underprediction was observed, as only low numbers are being predicted well. In a biological context, this means the model is able to predict only when there is not a *Microtus arvalis* outbreak.

Examining the individual relationships between the VAI and predictors in the model reveals a notable pattern with temperature (Figure 3A). There is a generally negative correlation up to just over 10°C, followed by a positive relationship that continues until around 15°C, after which it levels off slightly. Temperature has been linked with *Microtus arvalis* outbreak previously, for example, Imholt et al. (2011) investigated the role of weather factors in explaining regional differences in the risk of *Microtus arvalis* outbreaks in Eastern Germany, finding a strong link between weather conditions and outbreak risk across various locations and years. They suggested that weather data, when combined with other environmental factors like soil properties, could be used to develop a predictive outbreak model. Expanding on this in 2014, Esther et al., also in Germany, concluded that weather conditions were reliable predictors of *Microtus arvalis* outbreaks, emphasised more so than other factors including farming practices, soil properties, or landscape complexity. They observed that variations in vole population sizes correlated with specific weather conditions and their thresholds, and proposed that these relationships could be used to improve predictive models on a continental, and possibly global scale. Both studies (Imholt et al., 2011, Esther et al., 2014) therefore concluded that a well-developed model, based on a large dataset from Europe and incorporating weather and other environmental factors, could effectively predict *Microtus arvalis* outbreaks. However, despite accounting for these factors, and presenting a similar relationship with temperature, the present study was unable to predict *Microtus arvalis* outbreaks in Castilla y León. This contradictive result could potentially mean there should be further research into the confirmation of the previous studies' results, before further modelling using their conclusions. Alternatively, perhaps modelling techniques could be less transferrable between countries than previously thought.

Further inspection of the individual relationships the model has produced shows *Microtus arvalis* growth has a positive relationship with wooded pasture, showing a contrast to the negative relationship with tree pasture (Figures 3B and 3C), which potentially would have been expected to show similar relationships. One explanation for this could be related to the heterogeneity of the landscape provided by the trees. Structural elements in landscapes have been found to decrease population growth and outbreaks (Fotheringham et al., 2024). This is by providing habitat for predators and creating physical barriers that reduce connectivity between vole populations (Fotheringham et al., 2024). In this case, trees could be acting as these structural elements, which could explain the negative relationship between tree pasture and *Microtus arvalis* population growth. On the other hand, the model shows a particularly strong positive relationship with woody pasture, which contrasts with the negative relationship seen with tree pasture. This discrepancy might be due to the higher levels of biodiversity found in wooded pastures (Jaiteh, 2012), which could provide a greater variety of food resources and burrowing spaces for voles. Alternatively, these areas might be less intensively managed, allowing voles to thrive in conditions that are more favourable for their population growth.

The expected rodent population abundance relationship with soil type has also been studied in the past, with findings indicating soil texture does have an effect. Specifically, a preference was observed for sandy soils over clay soils, which was attributed to the better burrowing ability of sand (Massawe et al., 2008). However, in the present study, there is no apparent difference between the *Microtus arvalis* growth relationship with sand and clay soil, with both only having a subtle negative relationship (Figures 3D and 3E). Massawe et al. (2008) also noted that precipitation can negatively influence the burrowing ability of clay soils, consequently affecting the relationship of soil and rodent population numbers. Importantly, they found no significant difference between different soil types in dry seasons. Therefore, one possible explanation for the divergence between the previous findings and this study is that the soil data used here was collected during a season with lower precipitation. Similarly, the differing results could be due to a lower overall level of rainfall in Castilla y León compared to the site of the previous research, providing a barrier for transferability of conclusions between locations. To narrow down potential explanations and analyse this relationship further in future models, soil data that varies temporally as well as spatially could be used, such as measurements taken from each time period. This would allow deeper insights into whether differing soil type influences *Microtus arvalis* population size.

A factor that may be impacting the model's performance is the prolonged outbreak observed in the data, around periods 23-32**,** which reaches into the test data (Figure 1C). The model performed ineffectively in predicting VAI during these periods, as indicated by high RMSE (Figure 6), suggesting that the extended outbreak is influencing the model's accuracy. Notably from rodent cycle analysis in the past, data from across Europe found that from 1995 to 2005, vole population dynamics were abnormal (Cornulier et al., 2013). It was found that during this time, the extent of cycles was limited, showing patterns unlike that of the rest of the data. As the decline pattern was consistent across ecosystems, it was hypothesised a large-scale environmental factor was at fault – likely regarding climate change, although specific mechanisms could not be identified (Cornulier et al., 2013). Putting this into context with this study, the prolonged outbreak seen here is also an abnormal pattern. Arguably, this could be another unusual duration of amended cyclic pattern which is difficult to pinpoint the cause of. If this is the case, this was likely beyond the predictive capabilities of the model. If these irregular phases are linked with climate change as suggested, this means complex population dynamics such as rodent outbreaks and cycles will continue to become less predictable using quantitative methods.

As the model is not working as a predictor of *Microtus arvalis* outbreaks, future research could potentially benefit from using predictor variables beyond those which were used in this study. Or alternatively, adjustments could be made to the methods. Notably from the distribution of the VAI (Figure 1A), it consists largely of zeros, which is potentially at fault for the models underprediction. Arguably, alternative modelling techniques could have better allowed for this.

One way of achieving this would be using a zero-inflated model, allowing better for higher zero distributions within data (Dénes et al., 2015). For example, a ZIP (Zero-Inflated Poisson) model would differently account for the high number of zeros, (Dénes et al., 2015) potentially improving the predictions of a model with a high number of them, such as the one in this study. Another potential modification to improve the model could involve aggregating the response variable to simplify the prediction process (Pollet et al., 2015). This would involve converting the VAI into a binary format. For example, if an outbreak is defined as a VAI exceeding 0.5, the model's predictions could be shown as either outbreak or no outbreak. This approach could potentially increase predictive accuracy, as the model would only need to distinguish between two possible outcomes, which is far simpler than the multiple possibilities considered in this study. This method would further the importance of predicting far in advance, as rising numbers towards an outbreak would no longer be an indicator of an impending outbreak. Notably, this aggregation would also result in drawbacks for the model as the predictions would not be precise as attempted in this study (Pollet et al., 2015); further, a specific threshold that constituted an outbreak would be hard to define and may not be equally applicable in different cases.

If future studies are able to build on the foundational methods trialled in this paper, it would provide a basis for management strategies to be able to be aimed at predicting rather than reacting to the outbreaks. The model would provide information on when and where the outbreak will take place and predictors with strong relationships could further be used as indicators for when to implement management strategies. This would then allow for avoiding extreme management attempts and using more preventative measures before the outbreak peaks. Current proposed preventative measures often aim to disrupt habitats and breeding sites. One approach involves dragging soil along ditches and track edges (Rodríguez-Pastor et al., 2016), though this could also affect non-target species. Another method is ploughing or using less invasive tillage techniques (Roos et al., 2019), however any tillage can encourage climate change. A further strategy is promoting natural predators by installing perches and nest boxes to support kestrels and barn owls (Paz et al., 2013), but this biological suppression of *Microtus arvalis* is an area of little research (Machar et al., 2017), and the scale of outbreaks means it is potentially unlikely predators could control the population. Beyond these preventative actions, less extensive rodenticide could also be used (Lauret et al., 2019), yet this too poses risks to non-target species and is further complicated by EU regulatory restrictions. The potential issues with these approaches highlight effective control measures as an area still requiring thorough research. The ideal outline of future management would be methods which are not harmful to the surrounding environment and organisms, are inexpensive, effective and take into account social standpoints of the area as much as possible. Arguably, one method of control that could be promising is ribonucleic acid interference (RNAi), in which genes which are critical to survival are targeted to go through sequence-specific gene silencing (Campbell et al., 2015). This method is currently being largely developed towards agricultural insect pests, however RNAi, which in this case would be orally administered, has also succeeded in being lethal to mice (Martin et al., 2011). If this information could be used to further develop an RNAi based rodenticide for Microtus arvalis, it could potentially be an effective control measure. This rodenticide would not pose a risk to the environment, as well as being species-specific (Xue et al., 2012), avoiding harm to non-target species. Beyond developing the approach, future research would need to assess any potential risks this could pose and attempt to gauge potential social controversy before implementation.

Conclusions

This study developed a modelling approach to predict *Microtus arvalis* population fluctuations, with a view to providing a method to assist in implementing rodent outbreak management strategies before the outbreak causes extensive damage. The model predicted VAI more accurately in some time periods than others (predicting 16-28 months in advance), however they were the periods of lowest population numbers, showing the model can only predict when there is not an outbreak. This result was backed up by the model accuracy (RMSE) being calculated as higher than the mean of the response variable. Nevertheless, this study tested foundational methods which could be built on by future studies as well as giving potential insight into the predictive value of environmental characteristics used as predictors – for example temperature and soil type were not found to have predictive power, contradicting previous research. Future research can build on these conclusions by including predictors beyond those used here to test for more predictive power, or alternatively can make adjustments to the methods carried out to allow for predictions of outbreaks. If these adjustments were successful, these predictions would allow for management plans to be carried out earlier, minimising the extreme damage caused by rodent outbreaks.

Bibliography

Andreassen, H.P., Sundell, J., Ecke, F., Halle, S., Haapakoski, M., Henttonen, H., Huitu, O., Jacob, J., Johnsen, K., Koskela, E., Luque-Larena, J.J., Lecomte, N., Leirs, H., Mariën, J., Neby, M., Rätti, O., Sievert, T., Singleton, G.R., van Cann, J. and Vanden Broecke, B. (2020). Population cycles and outbreaks of small rodents: ten essential questions we still need to solve. *Oecologia*, 195(3), pp.601–622. doi:https://doi.org/10.1007/s00442-020-04810-w.

Aplin KP, Brown PR, Singleton GR, Douang Boupha B, Khamphoukeo K (2006). Rodent ecology in the rice environments of Laos. In: Schiller JM, Channengxay MB, Linquist B, Appa Rao S, eds. Rice in Laos. *International Rice Research Institute*, pp. 291-308.

Beckage, B., Gross, L.J. and Kauffman, S. (2011) 'The limits to prediction in Ecological Systems', *Ecosphere*, 2(11). doi:10.1890/es11-00211.1.

Blank, B., Jacob, J., Petri, A. and Esther, A. (2011). Topography and soil properties contribute to regional outbreak risk variability of common voles (Microtus arvalis). *Wildlife Research*, 38(7), pp.541–541. doi:https://doi.org/10.1071/wr10192.

Campbell, K.J., Beek, J., Eason, C.T., Glen, A.S., Godwin, J., Gould, F., Holmes, N.D., Howald, G.R., Madden, F.M., Ponder, J.B., Threadgill, D.W., Wegmann, A.S., Baxter, G.S. (2015) 'The next generation of rodent eradications: Innovative technologies and tools to improve species specificity and increase their feasibility on islands', *Biological Conservation*, 185, pp. 47–58. doi:10.1016/j.biocon.2014.10.016.

Constantino Caminero-Saldaña, Jennifer Paola Correa-Cuadros, Baños-Herrero, A., Riquelme, C., Yesica Pallavicini, Fernández-Villán, M., Plaza, J., Rodrigo Pérez Sánchez, Sánchez, N., François Mougeot, Juan José Luque-Larena, Jaksic, F.M. and María Carmen García-Ariza

(2023). Exploring the influence of density-dependence and weather on the spatial and temporal variation in common vole (*Microtus arvalis*) abundance in Castilla y León, NW Spain. *Pest Management Science*. doi:https://doi.org/10.1002/ps.7954.

Cornulier, T., Yoccoz, N.G., Bretagnolle, V., Brommer, J.E., Butet, A., Ecke, F., Elston, D.A., Framstad, E., Henttonen, H., Hörnfeldt, B., Huitu, O., Imholt, C., Ims, R.A., Jacob, J., Jędrzejewska, B., Millon, A., Petty, S.J., Pietiäinen, H., Tkadlec, E. and Zub, K. (2013). Europe-Wide Dampening of Population Cycles in Keystone Herbivores. *Science*, [online] 340(6128), pp.63–66. doi:https://doi.org/10.1126/science.1228992.

Daketse, M.-J. and Martinet, L. (1977). Effect of temperature on the growth and fertility of the field-vole, Microtus arvalis, raised in different daylength and feeding conditions. *Annales de Biologie Animale Biochimie Biophysique*, [online] 17(5A), pp.713–721. doi:https://doi.org/10.1051/rnd:19770607.

Dénes, F.V., Silveira, L.F. and Beissinger, S.R. (2015) ‘Estimating abundance of unmarked animal populations: Accounting for imperfect detection and other sources of zero inflation’, *Methods in Ecology and Evolution*, 6(5), pp. 543–556. doi:10.1111/2041-210x.12333.

Esther, A., Imholt, C., Perner, J., Schumacher, J. and Jacob, J. (2014). Correlations between weather conditions and common vole (Microtus arvalis) densities identified by regression tree analysis. *Basic and Applied Ecology*, 15(1), pp.75–84. doi:https://doi.org/10.1016/j.baae.2013.11.003.

Fotheringham, J., Rijsdijk, K. and Artzy-Randrup, Y. (2024). How landscape heterogeneity influences the outbreak dynamics of the common vole (Microtus arvalis) in European farmlands: A review. [online] doi:https://doi.org/10.3897/arphapreprints.e139043.

Garrett Grolemund, Hadley Wickham (2011). Dates and Times Made Easy with lubridate. *Journal of Statistical Software*, 40(3), 1-25. URL https://www.jstatsoft.org/v40/i03/.

Heroldová, M., Šipoš, J., Suchomel, J. and Zejda, J. (2021). Influence of crop type on common vole abundance in Central European agroecosystems. *Agriculture, Ecosystems & Environment*, 315, p.107443. doi:https://doi.org/10.1016/j.agee.2021.107443.

Herrero-Cófreces, S., François Mougeot, Lambin, X. and Juan José Luque-Larena (2021). Linking Zoonosis Emergence to Farmland Invasion by Fluctuating Herbivores: Common Vole Populations and Tularemia Outbreaks in NW Spain. *Frontiers in Veterinary Science*, 8. doi:https://doi.org/10.3389/fvets.2021.698454.

Högstedt, G., Seldal, T. and Breistøl, A. (2005). Period Length In Cyclic Animal Populations. *Ecology*, 86(2), pp.373–378. doi:https://doi.org/10.1890/02-0561.

Imholt, C., Esther, A., Perner, J. and Jacob, J. (2011). Identification of weather parameters related to regional population outbreak risk of common voles (Microtus arvalis) in Eastern Germany. *Wildlife Research*, 38(7), pp.551–551. doi:https://doi.org/10.1071/wr10196.

Jaiteh, K. (2012). *Wood Wise Woodland Conservation News Wood Pasture: definition, restoration, creation Ancient oak in Moccas Park NNR*. Available at: https://www.woodlandtrust.org.uk/media/1825/wood-wise-wood-pasture.pdf.

Jareno, D., Vi.uela, J., Luque-Larena, J.J., Arroyo, L., Arroyo, B. and Mougeot, F. (2015). Factors associated with the colonization of agricultural areas by common voles Microtus arvalis in NW Spain. Biological Invasions, 17(8), pp.2315–2327. doi:https://doi.org/10.1007/s10530-015-0877-4.

JCYL (Junta de Castilla y León). 2008. www.jcyl.es/scsiau/Satellite/up/es/AgriculturaGanaderia/Page/PlantillaDetalleContenido/1141143048675/Comunicacion/1213708951343/?asm=jcyl.

Krebs, C.J. and Myers, J.H. (1974). Population Cycles in Small Mammals. [online] *Academic Press*, pp.267–399. doi:https://doi.org/10.1016/S0065-2504(08)60280-9.

Lambin, X., Bretagnolle, V. and Yoccoz, N.G. (2006). Vole population cycles in northern and southern Europe: Is there a need for different explanations for single pattern? *Journal of Animal Ecology*, 75(2), pp.340–349. doi:https://doi.org/10.1111/j.1365-2656.2006.01051.x.

Lauret, V., Delibes-Mateos, M., Mougeot, F. and Arroyo-Lopez, B. (2019). Understanding conservation conflicts associated with rodent outbreaks in farmland areas. *Ambio*, 49(5), pp.1122–1133. doi:https://doi.org/10.1007/s13280-019-01256-0.

Li, G., Yin, B., Li, J., Wang, J., Wei, W., Bolnick, D.I., Wan, X., Zhu, B. and Zhang, Z. (2020). Host-microbiota interaction helps to explain the bottom-up effects of climate change on a small rodent species. *The ISME Journal*, 14(7), pp.1795–1808. doi:https://doi.org/10.1038/s41396-020-0646-y.

Luque-Larena, J.J., Mougeot, F., Roig, D.V., Lambin, X., Rodríguez-Pastor, R., Rodríguez-Valín, E., Anda, P. and Escudero, R. (2015). Tularemia Outbreaks and Common Vole (Microtus arvalis) Irruptive Population Dynamics in Northwestern Spain, 1997–2014. *Vector-Borne and Zoonotic Diseases*, 15(9), pp.568–570. doi:https://doi.org/10.1089/vbz.2015.1770.

Luque-Larena, J.J., Mougeot, F., Vi.uela, J., Jare.o, D., Arroyo, L., Lambin, X. and Arroyo, B. (2013). Recent large-scale range expansion and outbreaks of the common vole (Microtus arvalis) in

NW Spain. *Basic and Applied Ecology*, 14(5), pp.432–441. doi:https://doi.org/10.1016/j.baae.2013.04.006.

Machar, I., Harmacek, J., Vrublova, K., Filippovova, J. and Brus, J. (2017) 'Biocontrol of common vole populations by avian predators *versus* rodenticide application', *Polish Journal of Ecology*, 65(3), pp. 434–444. doi:10.3161/15052249pje2017.65.3.010.

Martin, J.N., Wolken, N., Brown, T., Dauer, W.T., Ehrlich, M. E., Gonzalez-Alegre, P. (2011) 'Lethal toxicity caused by expression of Shrna in the mouse striatum: Implications for therapeutic design', *Gene Therapy*, 18(7), pp. 666–673. doi:10.1038/gt.2011.10.

Massawe, A.W., Rwamugira, W., Leirs, H., Makundi, R.H., Mulungu, L., Ngowo, V. and Machang'u, R. (2008). Soil type limits population abundance of rodents in crop fields: case study of the multimammate rat Mastomys natalensis Smith, 1834 in Tanzania. *Integrative Zoology*, 3(1), pp.27–30. doi:https://doi.org/10.1111/j.1749-4877.2008.00070.x.

Olea, P.P., I.S. SáNchez-Barbudo, J. ViñUela, I. Barja, P. Mateo-TomáS, A. PiñEiro, R. Mateo, and F.J. Purroy. 2009. Lack of scientific evidence and precautionary principle in massive release of rodenticides threatens biodiversity: Old lessons need new reflections. *Environmental Conservation* 36: 1. https://doi.org/10.1017/S0376892909005323.

Pollet, T.V., Stulp, G., Henzi, S.P., Barrett, L. (2015) 'Taking the aggravation out of data aggregation: A conceptual guide to dealing with statistical issues related to the pooling of individual-level observational data', *American Journal of Primatology*, 77(7), pp. 727–740. doi:10.1002/ajp.22405.

R Core Team (2023). _R: A Language and Environment for Statistical Computing_. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.

Rodríguez-Pastor, R., Escudero, R., Vidal, D., François Mougeot, Arroyo, B., Lambin, X., Ave Maria Vila-Coro, Rodríguez-Moreno, I., Anda, P. and Juán José Luque-Larena (2017). Density-Dependent Prevalence of *Francisella tularensis* in Fluctuating Vole Populations, Northwestern Spain. *Emerging Infectious Diseases*, 23(8), pp.1377–1379. doi:https://doi.org/10.3201/eid2308.161194.

Rodríguez-Pastor, R., Luque-Larena, J.J., Lambin, X., Mougeot, F. (2016) '"Living on the edge": The role of field margins for common vole (Microtus arvalis) populations in recently colonised Mediterranean farmland', *Agriculture, Ecosystems & Environment*, 231, pp. 206–217. doi:10.1016/j.agee.2016.06.041.

Roos, D., Caminero Saldaña, C., Arroyo, B., Mougeot, F., Luque-Larena, J.J. and Lambin, X. (2019). Unintentional effects of environmentally-friendly farming practices: Arising conflicts between zero-tillage and a crop pest, the common vole (Microtus arvalis). *Agriculture, Ecosystems & Environment*, 272, pp.105–113. doi:https://doi.org/10.1016/j.agee.2018.11.013.

Roos, D. Caminero Saldaña, C., Elston, D., Mougeot, F., Garcia-Ariza, M.C., Arroyo, B., Luque-Larena, J.J., Revilla, F.J.R., Lambin, X. (2022) *From pattern to process? dual travelling waves, with contrasting propagation speeds, best describe a self-organised spatio-temporal pattern in population growth of a cyclic rodent, Ecology Letters,* 25(9), pp.1986-1998. doi:10.22541/au.164302552.25540752/v1.

Rouco, C., Juan José Luque-Larena, Vidal, D. and François Mougeot (2023). Vole outbreaks may induce a tularemia disease pit that prevents Iberian hare population recovery in NW Spain. *Scientific Reports*, [online] 13(1). doi:https://doi.org/10.1038/s41598-023-30651-7.

Simpson G (2024). _gratia: Graceful ggplot-Based Graphics and Other Functions for GAMs Fitted using mgcv_. R package version 0.9.2, <https://gavinsimpson.github.io/gratia/>.

Singleton, G. (2003). *Impacts of rodents on rice production in Asia*. [online] Available at: https://www.semanticscholar.org/paper/Impacts-of-rodents-on-rice-production-in-Asia-Singleton/fe942173ce1a17769a838f67430a8e4add485b11 [Accessed 20 Nov. 2024].

Singleton, G.R. and International Rice Research Institute (2010). *Rodent outbreaks : ecology and impacts*. Page 207-217. Metro Manila, Philippines: Irri.

Singleton, G.R., Brown, P.R., Jacob, J. and Aplin, K.P. (2007). Unwanted and unintended effects of culling: A case for ecologically-based rodent management. *Integrative Zoology*, 2(4), pp.247–259. doi:https://doi.org/10.1111/j.1749-4877.2007.00067.x.

Watts, J. (2009). *Chinese try to curb 'plague of desert rats' in Tibet with contraceptives*. [online*] the Guardian.* Available at: https://www.theguardian.com/environment/2009/mar/25/china-gerbils-deserts [Accessed 21 Nov. 2024].

Wickham, H. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York, 2016.

Wood, S.N. (2011) Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society* (B) 73(1):3-36

Xue-Yi, X., Ying-Bo, M., Xiao-Yuan, T., Yong-Ping, H., Xiao-Ya, C. (2012) 'New approaches to agricultural insect pest control based on RNA interference', *Advances in Insect Physiology*, pp. 73–117. doi:10.1016/b978-0-12-387680-5.00003-3.