The last wild grazer: Exploring habitat preferences of *Oryctolagus cuniculus* in the Maltese garrigue

Jérémie Tranchant

A dissertation presented to the Institute of Earth Systems of the University of Malta for the degree of Master of Science in Rural & Environmental Sciences

March 2024

L-Università ta' Malta

FACULTY/INSTITUTE/CENTRE/SCHOOL Institute of Earth Systems

DECLARATIONS BY POSTGRADUATE STUDENTS

(a) Authenticity of Dissertation

I hereby declare that I am the legitimate author of this Dissertation and that it is my original work.

No portion of this work has been submitted in support of an application for another degree or qualification of this or any other university or institution of higher education.

I hold the University of Malta harmless against any third party claims with regard to copyright violation, breach of confidentiality, defamation and any other third party right infringement.

(b) Research Code of Practice and Ethics Review Procedures

I declare that I have abided by the University's Research Ethics Review Procedures. Research Ethics & Data Protection form code IES-2022-00029.

As a Master's student, as per Regulation 77 of the General Regulations for University Postgraduate Awards 2021, I accept that should my dissertation be awarded a Grade A, it will be made publicly available on the University of Malta Institutional Repository.

02.2023

ABSTRACT

The European rabbit (*Oryctolagus cuniculus*, Linnaeus 1758) is the only remaining wild mammal herbivore in Malta, where it has been introduced during Antiquity. The species poses a global conservation paradox, being endangered in its native range but highly invasive elsewhere, in particular on islands. While Malta shares some characteristics with its native habitats, the ecology of the rabbit has never been assessed locally. This study aimed to assess the reality of this paradox in the Maltese context by quantifying the effect of selected habitat parameters of known importance to the rabbit in its native semi-arid ecosystems. A total of 24 sites (1ha each) were surveyed by integrating high-resolution drone imagery and national LiDAR data through supervised SVM classification (kappa = 0.83), along systematic quadrat surveys to measure pellet density. Analysis was performed using separate Generalized Linear Models relating pellet counts with shrub vegetation's structure (number of shrubs, shrub height and area, aggregation), intra-site land cover composition, and the surface of three anthropogenic features in and around sites (paved roads, hiking trails and rubbles walls), respectively. The extent of paved roads and hiking trails were both negatively correlated to pellet counts and explained together 48% of the variations in pellets across sites. The land cover composition could explain only 26% of the variations in pellets, while the ratio of herbaceous to low-lying shrubs was identified as the main driver negatively correlated to pellet density (marginal R^2=0.24). Lastly, the measured shrub metrics did not show significant effects on pellet counts. The findings of this study confirmed the predominant role of human development in Malta for the spatial distribution of the wild rabbit in the remaining open semi-natural habitats. Moreover, it highlighted that human hunting pressure is likely stronger than non-human predation pressure locally. Finally, a key role of the rabbit in the competition between herbaceous and dwarf shrub species in the Maltese garrigue was suggested. These findings have the potential to inform both ecological restoration projects and wild rabbit management in Malta.

Keywords: Rabbit Relative Density Index, Habitat Preferences, Unmanned Aerial Systems (UAS), Food and Refuge Availability, SVM classification, Land Use Land Covers, Photogrammetry.

Table of Contents

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AIC: Alkaike Information Criterion
AICc: Alkaike Information Criterion Corrected
AGM: Above Ground Model
ANN: Average Nearest Neighbour
AOI: Area of Interest
BIC: Bayesian Information Criterion
C2: Command & Control
CORS: Continuously Operating Reference Station
DEM: Digital Elevation Model
DPC: Dense Point Cloud
DMF: Depth Maps Filtering
DSM: Digital Surface Model
DTM: Digital Terrain Model
DV: Dependent Variable
EF: Exponential Family
GCP: Ground Control Point
GCS: Ground Control Station
GDS: Ground-Sampling Distance
GIS: Geographic Information System
GLM: Generalized Linear Model
GNSS: Global Navigation Satellite System
GSD: Ground-Sampling Distance
HSV: Hue-Saturation-Value
HNS: Hybrid Navigation System
ICAO: International Civil Aviation Organization
IDH: Intermediate Disturbance Hypothesis
IUCN: International Union for Conservation of Nature
LiDAR: Light Detection and Ranging
LR: Log-Ratio
ML: Maximum Likelihood
MLE: Maximum Likelihood Estimation
MSL: Milchunas–Sala–Lauenroth
MVS: Multi-View Stereo
NB: Negative Binomial
NN: Nearest Neighbour
PA: Planning Authority (of Malta)
RF: Random Forest
RGB: Red-Green-Blue
RHD: Rabbit Haemorrhagic Disease
RMSE: Root-Mean Standard Error
RTK: Real-Time Kinematics
SfM: Structure-from-Motion
SPC: Sparse Point Cloud
STP: Standing crop pellets
SVM: Support Vector Machine
UAS: Unmanned Aerial System
UAV: Unmanned Aerial Vehicle
US: Unmanned System

ACKNOWLEDGEMENTS

This Master's Dissertation owes its completion to the generous support and guidance of numerous individuals and organizations, for which I am profoundly grateful. Foremost, I extend my sincere thanks to my supervisors, Dr Belinda Gambin and Prof Sandro Lanfranco, whose expertise and encouragement were instrumental in my research journey. The University of Malta and the Ministry of Transport, Infrastructure, and Planning of Malta generously provided essential equipment. I am also grateful for the unwavering support and advice from the entire Research and Planning Unit at MTIP, with a special mention of Daniel Fenech, Christopher Gauci and Emanuele Colica for their technical guidance and Dr George Buhagiar for his . I am thankful to Dr Owain Barton and Tim van Berkel for their statistical guidance on modelling, along with Dr Fiona Sammut from the Department of Statistics for her insightful advice on compositional analysis. My gratitude also goes to Arthur Lamolière and the CLAS for granting access to their processing equipment. Additionally, I am appreciative of the Ranger unit of Il-Majjistral Park and Alex Casha for providing relevant literature, and to Jerôme Letty from the Office Français de la Biodiversité (OFB) for his guidance on rabbit density. Finally, the unwavering support of my family and my partner Naomi Ruiz carried me through this five-years long professional reconversion.

1. Introduction

1.1 Research context

Oryctolagus cuniculus is an herbivorous mammal native to the Iberian Peninsula which has been dispersed by humans to all continents and more than 800 islands during the last millennia (Flux & Fullagar 1992). Outside its natural range, it is generally considered an agricultural pest and is the target of numerous control or eradication programs, particularly on islands (Courchamp, Chapuis & Pascal 2003). In its natural European range, on the other hand, its population experienced a sharp decline during the 21st century, leading to its recent classification as Endangered by the IUCN Red List (Villafuerte & Delibes-Mateos 2019). Its demographic decline in Europe resulted in local trophic cascades and the recognition of the European rabbit as a keystone species in its native range, placing it at the centre of a renewed conservation effort in Spain, Portugal, and France (Delibes-Mateos et al. 2007).

This ambivalence has led some researchers to consider the European rabbit as a conservation paradox (Delibes-Mateos et al. 2011; Lees & Bell 2008). The three countries mentioned above subsidize both control and conservation measures implemented inside territories that are close enough to influence each other, providing mixed outcomes (Delibes-Mateos et al. 2011). The

effects of the European rabbit on trophic webs are complex and driven by the local assemblages of native and exotic species (Barbar & Lambertucci 2018; Lees & Bell 2008). The impacts of its herbivory can shape entire communities, due to its high reproduction rate and its small size, which allows higher densities and complicates control procedures compared to larger mammal herbivores. Where the rabbit has been introduced, its eradication does not always lead to community recovery due to the rapid ecological adaptations often following introduction (Lees & Bell 2008). On the other hand, in the Mediterranean ecosystems with which the rabbit has evolved, it often performs several ecosystem functions by contributing to seed dispersal, increasing floral and faunal diversity, changing soil fertility and, above all, playing a central prey role in trophic webs (Delibes-Mateos et al. 2007; Eldridge et al. 2010, Galvez-Bravo et al. 2008; Lees & Bell 2008; Willott et al. 2000).

In the Maltese archipelago, wild European rabbits were supposedly introduced during antiquity and represent an important element of the Maltese culture under various aspects (Cassar 1994; Hardy et al. 1994). Today, it is the only wild terrestrial mammal herbivore of the Maltese archipelago and preyed upon by only two wild mammal predator species, the Least Weasel and the Ferret, along with the domestic cat. This absence of significant terrestrial predators minimizes the trophic role of the European rabbit in Malta; however, the archipelago is an important stepping stone for several migratory raptor species, which are also often hunted with known issues (Falzon 2008).

The Maltese ecosystems belong mostly to the Western Mediterranean ecological succession (Vogiatzakis & Griffiths 2008) and share several plant species with the Iberian Peninsula, such as the Mediterranean thyme (*Thymbra capitata*), also known as the Spanish oregano, or the Mediterranean Heath (*Erica multiflora*). There are also various relict populations inherited from ancient connections with North Africa and Sicily, resulting in communities with a high level of

endemism and species richness (Cassar, Conrad & Schembri 2008; Hunt & Schembri 1999). It is likely that the plant communities have adapted to grazing given the long-standing free-range husbandry practices in Malta (Farrell, Hunt & McClung 2020; Schembri & Lanfranco 1993). However, these practices gradually disappeared during the 20th century (Attard 2008), leaving the wild European rabbit as the sole mammal grazer in the Maltese landscapes. In this context, the ecological associations between *Oryctolagus cuniculus* and Maltese ecosystems are unclear, with potential both for invasiveness due to the insular setting and keystone functions due to local adaptations to grazing and seasonal predation.

The species' current legal status somewhat reflects the conservation paradox mentioned above, because it is traditionally considered and managed as an agricultural pest while at the same time being subject to protection regulations as a game species, with some areas where hunting is banned entirely (Protection of Wild Rabbit Regulation, SL 549.90 of 2014). Local wild populations have dwindled during the 21st century due to viral episodes and anthropic pressures (Muscat 2021), however, quantified scientific assessments are absent. In fact, the current distribution and abundance of the local wild European rabbit population and its relationships to the Maltese ecosystems are still largely unknown. Scientific initiatives have recently been undertaken to fill this gap, with an on-going effort focused on population inventories and the establishment of methods adapted to the local context the Wild Rabbit Research Initiative.

The spatial distribution of *Oryctolagus cuniculus* is determined through indirect (i.e. observing signs of presence such as latrines, dung pellets, burrows, etc…) or direct (i.e. direct sighting counts) abundance measurements. These methods are technically simple but costly in human resources to cover large areas. In Malta, a citizen science application has been deployed in the agricultural community to report sightings of wild rabbits, in order to map their distribution at

coarse spatial resolution (Mr. I. Falzon 2022, personal communication, 11 August). Another approach through modelling could allow to predict the presence and density of European rabbit populations through various ecological associations that are often specific to the local environment (Fernandes 2012; Kontsiosis et al. 2013).

While the European rabbit is known for its ecological plasticity, it has been shown that plant communities influence the distribution of *Oryctolagus cuniculus* in Mediterranean ecosystems (Monzon, Fernandes & Rodrigues 2004). In particular, the presence of shrubs and the degree of fragmentation of the landscape seem to be important elements with contrasting effects on rabbit density, reinforcing the need for local assessment (Dellafiore, Gallego Fernandez & Munoz Vallés 2008; Monzon, Fernandes & Rodrigues 2004). Wild rabbits use shrubs and rocky interstices as cover for warrens and as temporary refuges during feeding or dispersal periods, often favouring landscapes that contain semi-open scrublands in the vicinity to richer feeding grounds (Dellafiore, Gallego Fernandez & Munoz Vallés 2008). The species of shrubs and their morphologies have also been found to be determinants of the distribution of European rabbits inside Mediterranean scrublands, because they tend to prefer shrubs offering good overhead cover but low horizontal cover to facilitate the detection of ground predators (Beja, Pais & Palma 2007). On the other hand, wild rabbits feed mainly on grasses and forbs, even though they can adapt their feeding habits to a wide range of sources, including woody parts and mildly toxic plants (Kontsiosis et al. 2018). They tend to favour landscapes offering a balance between refuge and food availability, but in their native range where predation pressure is significant, they prefer a higher refuge potential over food availability and extend their home range to forage further (Carvalho & Gomes 2004; Martins et al. 2003). Lastly, soil suitability for digging has been found to be a significant determinant of rabbit presence, often measured through soil hardness, soil depth or bedrock cover (Encarnação 2018).

Recent advances in Unmanned Aerial Systems (UAS) have allowed to derive precise land cover surfaces and various vegetation measurements, including above-ground elevation, at lower costs than traditional field surveys (Fraser et al. 2016; Howell et al. 2020; Perez-Luque et al. 2022; Ventura et al. 2022). In Malta, few studies have assessed the potential of small UAS for land cover assessment (Bellia & Lanfranco 2019; Zerafa et al. 2023), but there is a growing interest for the applications of this technology accompanied by various equipment acquisitions (Gatt 2021). Until now, there has been few attempts to use UAS in applied conservation biology locally. The Maltese environment poses various challenges due to scale and landscape heterogeneity which provide stimulating opportunities for UAS research.

1.2 Justification of the research

Knowledge on the associations between vegetation structure, land cover and wild European rabbit density would improve both population and landscape management in Malta, depending on conservation objectives of local rabbit populations or on the other hand, localized landscape restoration objectives. The definition of broad habitat preferences of the Maltese wild European rabbit could also facilitate national-scale distribution mapping through predictive modelling using remotely-acquired imagery. Lastly, it may provide insights into the insular ecology of wild *Oryctolagus cuniculus* in a context of low animal predation, high habitat fragmentation and strongly anthropogenic landscapes.

This research aims to explore the influence of food and refuge availability on wild European rabbit density in the Maltese garrigue through relative density index and remote-sensing. The dissertation asks two main research questions and a third auxiliary research question:

I. Is the relative rabbit density influenced by the characteristics of the shrub vegetation in the Maltese garrigue?
    - H0: There is no relationship between shrubs structure and relative rabbit density in the Maltese garrigue.

II. Is the relative rabbit density influenced by land cover composition in the Maltese garrigue?
    - H0: There is no relationship between the land cover composition of the Maltese garrigue and relative rabbit density.

III. Is the relative rabbit density in the Maltese garrigue influenced by anthropogenic factors?
    - H0: There is no relationship between anthropogenic features and relative rabbit density in the Maltese garrigue.

1.3 Structure of the dissertation

The second chapter of this dissertation review the existing literature on the Maltese environment to provide information on the background to this study. The climatic, geologic, and ecologic contexts are presented with a focus on the garrigue community, and ecological theories on disturbance are put in perspective with the Maltese history of grazing. In a second part, the literature on the wild European rabbit biology and ecology is reviewed along with the techniques used to assess its abundance and the local Maltese context. A third section reviews the theory behind the use of UAS and photogrammetry for terrain modelling, landscape surveys and imagery classification. The two last sections present the theoretical background for statistical modelling of compositional and count data respectively.

The third chapter of the dissertation presents the methodology used for data collection and analysis. The methods used for the relative rabbit density survey and for the drone survey of the landscape are presented separately.

The fourth chapter presents the results of this dissertation. The accuracy assessments of the photogrammetry and imagery classification processes are first reported, followed by the analysis of the dung pellets data representing the dependent variable used in this dissertation. It is followed by descriptive statistics for the two sets of explanatory variables considered in this study: the shrubs metrics and the land cover composition. Lastly, the results of the modelling procedures are reported separately for each set of explanatory variables.

The fifth chapter is dedicated to the discussion of the reported results, followed by the conclusion of this dissertation. The cited literature can be found in the 'References' section while the 'Appendices' provides all the detailed datasets, programming code, tables and in-depth methods referred to in the text.

2. Literature review

2.1 The Maltese environment

2.1.1 Climate

The Maltese archipelago experiences a bi-seasonal climate from within the Sicily Channel in the Central Mediterranean Sea. The hot and dry summer, extending from April to September, contrasts with the mild and wet winter, lasting from October to March, which accounts for 85% of the annual precipitation totalling 530mm per year (Schembri 1997). With over 300 clear days annually, the archipelago had a mean monthly bright sunshine duration of 8.1 hours per day in 2020 and a mean cloud cover of 3.1 oktas between 1991-2020 (Galdies 2022). Over the same period, the average annual temperature was 19.4°C, while the relative humidity remained fairly consistent throughout the year, ranging between 65% and 80%. The dominant winds are from the northwest (the Mistral wind or Majjistral in Maltese) and the west, with an annual mean speed of 15.2 km/h and an annual mean gust speed of 71 km/h over the 1991-2020 period (Galdies 2022).

2.1.2 Geography, geology, and geomorphology

Characterized by its lack of mountains, the Maltese archipelago does not exhibit large-scale climatic, edaphic, or ecological gradients related to elevation (Cassar, Conrad & Schembri 2008). Comprised of three principal islands (Malta, Gozo, and Comino) and numerous islets, the archipelago formed due to extensional tectonics and uplift in the central regions of the Pelagian Block during the Late Miocene period, coinciding with the development of the Pantelleria Rift System (Chatzimpaloglou et al. 2020). The primary rock formations are generally horizontal, displaced by two types of faults: (i) east-northeast to west-southwest trending faults and (ii) northwest to southeast trending faults (the Magħlaq and related faults). These faults caused Malta's western side to be elevated compared to the east, leading to steep cliffs along the western and northwestern shores (Chatzimpaloglou et al. 2020). Furlani et al. (2018) noted that Malta has remained tectonically stable throughout the Holocene until today. The Maltese geology primarily consists of sedimentary marine limestone formations from the Oligo-Miocene period, approximately 30 to 5 million years ago, along with some Quaternary deposits (Schembri 1993). Malta's rock formations can be classified into five main types, arranged in descending order of age: Lower Coralline Limestone, Globigerina Limestone, Blue Clay, Greensand, and Upper Coralline Limestone (Chatzimpaloglou et al. 2020; Schembri 1993).

This geological and tectonic history shaped the geomorphology of the Maltese Islands and resulted in a notable northeast tilt (Cassar, Conrad & Schembri 2008; Chatzimpaloglou et al. 2020). It causes the Lower Coralline Limestone to be found along the western and southern steep cliffs of Malta, where the highest point of the islands, Ta' Zuta, lies at 253m above sea level (Figure 2.1). On the eastern and south-eastern coasts, the tilt results in gently sloping shores and drowned valleys forming a ria and bay coastline. The five main rock formations in

Malta are only present north of the Great Fault and on the Rabat-Dingli uplands south of it (Figure 2.1). Elsewhere, the Globigerina Limestone forms plains, shallow depressions, and low hills. The Upper Coralline Limestone forms karstic plateaux with steep cliffs and boulder screes at their base, formed by the erosion of underlying soft rock. This erosion also causes the Blue Clay to form taluses and the Upper Coralline Limestone boulders to travel down-slope, creating a boulder-strewn shoreline locally named *rdum*.

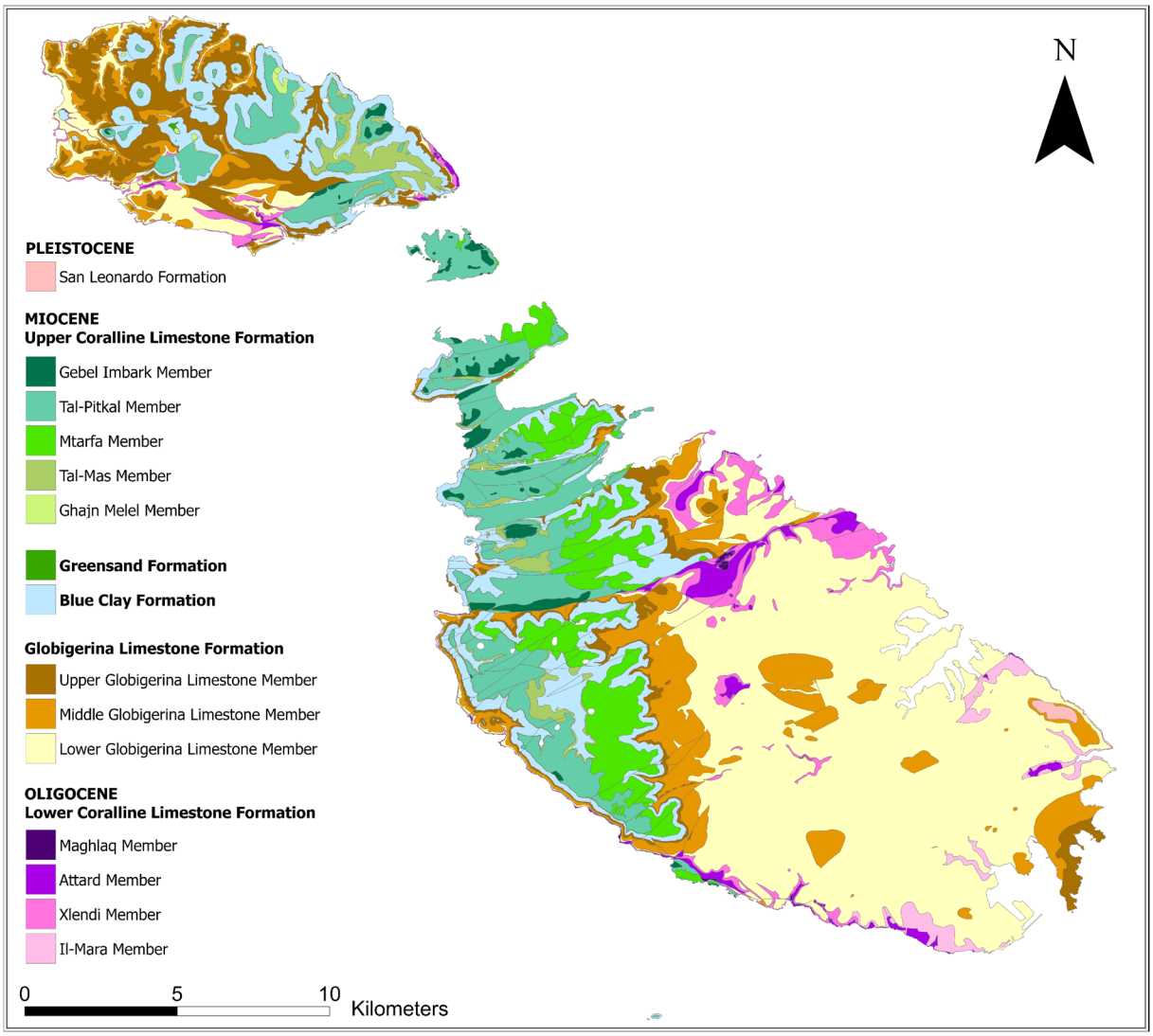


Figure 2.1. *Geological map of the Republic of Malta. Adapted from Continental Shelf Department, Malta, 2022.*

The local soils are generally characterized by their human alteration, their overall lack of horizon and general thinness. In 2013, the majority of Maltese soils had less than 2% of organic matter in their topsoil layer, the average soil depth was 48cm, the maximum soil depth was 200cm, and 23% of soils were less than 20cm deep, typically on plateaux and eroded surfaces often corresponding to garrigue biotopes (Sultana 2017). Maltese soils were divided by Lang (1960) according to the Kubiena classification system into three broad categories, with the "Terra soils" being a relic soil most commonly found on the karstic plateaux (Schembri 1993). A more recent classification system, MALSIS (2003), distinguished three main soil types associated with garrigues: (i) calcisols, showing a significant amount of translocated calcium carbonate, (ii) leptosols, shallow calcareous soils with a high gravel content, associated with low steppic and low garrigue vegetation, (ii) luvisols, reddish relict soils that formed under a previous wetter climatic regime, showing a small horizon and associated with high garrigue and maquis vegetation (Chatzimpaloglou et al. 2020). In addition, the MALSIS system identified several soil landscape types each containing "a range of soil types with a defined relationship within the landscape", including the Garrigue type (Vella 2005, p. 238).

2.1.3 Ecology

The Maltese archipelago has seen three main episodes of species influx due to marine regressions during the Pleistocene period, and a possible minor influence of tectonic activity (Cassar, Conrad & Schembri 2008; Schembri & Lanfranco 1996). They were each followed by prolonged periods of isolation, leading to the development of unique features such as gigantism, nanism and flightlessness in certain species. The presence of species with North African affinity on the islands could be due to their arrival by land during the Messinian marine regression, passive dispersal, or a combination of both, while the last influx corresponds to the arrival of humans to the islands at around 7500-7000 BP. Hunt and Schembri (1999, p.33)

summarized the Maltese biogeographic history as "a series of turnovers" with a biota replacing another after each large environmental change. Vogiatzakis and Griffiths (2008, p.68) mentioned that those discontinued connections with the surrounding continents have led the current biota of Malta to be part of the "Western Mediterranean biogeographical realm".

As stated by Schembri (1993), the majority of the current vegetation communities in Malta belong to the Mediterranean ecological succession sequence, which ranges from steppic grassland to phrygana, garrigue, maquis, and woodland, with the latter being the theoretical climax community. However, in Malta, the only fully functioning semi-natural forest ecosystem is found at Buskett (Schembri 1993). On the other hand, minor vegetation communities have been described as specialized, rare or relics communities from previous climatic regimes surviving in niches. Vogiatzakis & Griffiths (2008, p.66) summarized the definitions of each Mediterranean vegetation community type in Table 2.1 below.

Table 2.1 Main vegetation community types of the Mediterranean ecological succession sequence, from Vogiatzakis and Griffiths (2008, p.8).

Term	Definition
Forest	Tree-covered land but can also include other habitats in a matrix of trees
Maquis	A dense mostly evergreen shrub community 1–3 m high characteristic of the Mediterranean region
Garrigue	A community of low scattered often spiny and aromatic shrubs of the Mediterranean region
Phrygana	Low shrub developed over dry stony soil in the Mediterranean region. In general it is an equivalent term to garrigue which is used in the West Mediterranean.
Steppe	Composed of grasses, bulbous and other herbaceous plants
Savanna	A term that denotes big trees widely spaced with an understorey of phrygana or steppe (Grove and Rackham 2001)

2.1.3.1 Garrigue

Garrigue is the most common vegetation community of the Maltese islands, typically found on karstic limestone substrates with shallow soil, such as the plateaux of western Malta and the Gozitan hills (Chatzimpaloglou et al. 2020; ERA 2018; Schembri 1993; Schembri et al. 2009). However, the term 'garrigue' encompasses a wide range of phytosociological associations in Malta, which have been described by Brullo et al. (2020). Garrigues can result from naturally limiting abiotic factors (e.g. soil depth, wind) which prevent ecological succession, or from the degradation of forest and maquis (e.g. deforestation), often from anthropogenic activity. In the latter case, the garrigue is often surrounded by a more extensive habitat of a different type. It may be part of an impoverished maquis that results from the degeneration of climax woodland due to human activities, such as cutting and grazing, or it may be part of a steppic assemblage dominated by grasses, umbellifers, thistles, and geophytes following the degradation of the maquis and garrigue due to grazing and other human activities. Cassar, Conrad and Schembri (2008) state that the Maltese garrigue ecosystem encompasses various subtypes, each characterized by a dominant shrub species: Mediterranean Thyme (*Thymbra capitata*, sagħtar), the Maltese Yellow Kidney Vetch (*Anthyllis hermanniae*, ħatba sewda), Olive-leaved Germander (*Teucrium fructicans*), Mediterranean Heath (*Erica multiflora*), and the endemic Maltese Spurge (*Euphorbia melitensis*, tengħud tax-xagħri). Additionally, mixed garrigues featuring two or more of these dominant species are also prevalent, along with the presence of numerous geophytes and therophytes. The main phytosociological associations constitutive of Maltese garrigues are either from the *Cisto eriocephali-Ericion multiflorae* alliance, with the *Erico multiflorae-Antyllidetum melitensis* characterized by microphyllous xerophytes and the degraded maquis of *Periploco angustifoliae-Euphorbietum dendroidis*, or to the *Crithmo*

maritimi-Limonietea alliance when it is found closer to the coast as subhalophilous communities such as the *Anthyllido melitensis-Euphorbietum melitensis* (Brullo et al. 2020).

Steppic vegetation communities are often found along garrigues and include several grass species depending on the local topography and on the succession configuration of a given assemblage. Like garrigues, steppes also result either from naturally limiting environmental conditions or from the degradation of scrubland or maquis. In the first case, grass species tend to dominate the communities, whereas in the second case, selective grazing pressure favours adapted species such as thistles, rosette-forming or bulbous species (Farrell, Hunt & McClung 2020). The main phytosociological steppic associations found with garrigues in Malta are often the *Chamaeleo gummiferis-Brachypodietum retusi*, characterized by the False Brome (*Brachypodium retusum)*, the *Hyparrhenietum hirto-sinaicae* and the *Allietum lojaconoi* (Brullo et al. 2020). These communities contain species such as Esparto Grass (*Lygeum spartum*), Beard Grass (*Hyparrhenia hirta*), *Andropogon distachyus*, Canary Grass (*Phalaris truncata*), though the latter is rare, Steppe Grass (*Stipa capensis*), Goat Grass (*Aegilops geniculate)*, and various thistles, such as *Carlina involucrata*, *Notobasis syriaca*, and *Galactites tomentosa*. The common geophytes Branched Asphodel (*Asphodelus ramosus*) and Sea Squill (*Drimia maritima*) are commonly found in these assemblages, with a higher prevalence in degraded steppes (Brullo et al. 2020).

2.1.3.2 Shrubs

Shrub species are small to medium sized perennial woody plants, deciduous or evergreen, that have vital ecological functions in Mediterranean garrigue ecosystems (Lombardo et al. 2020). Their diverse adaptations to drought, fire and grazing increase their community's resilience to these pressures (Aronne & Wilcock 1994; Díaz Barradas et al. 1999; Hellicar & Kirschel 2022; Lombardo et al. 2020). The two major types of reproduction in Mediterranean shrubs, obligate resprouters and obligate seeders, allow overall continuous

shrub vegetation cover to be maintained in the face of severe disturbances such as fire (Keeney 1986). Obligate resprouters require specific microsites, such as deep soils or fissures, with underground water during summer droughts, have high seed dispersability and immediate establishment, but low growth rate in water-stressed conditions (Keeney 1986). Once in a favourable site, they can maintain themselves in it for a long time through replacement of the canopy and resprouting after fires or grazing. Obligate seeders, on the other hand, can establish on most substrates, have rapid growth rates, high drought tolerance, low shade tolerance, low seed dispersal and high seed longevity. They establish after fires and are maintained in gaps between sprouters. Obligate seeders' abundance increases as the frequency of safe sites for sprouters decreases, and once established, new sites are available only after disturbance, typically fire (Keeney 1986).

Shrub species shape diversity patterns in arid and semi-arid ecosystems by providing refuge (i.e., 'nurse species') for other plant species from abiotic (e.g., high aridity, fire) and biotic (e.g., overgrazing, competition) stress (Rahmanian et al. 2021). Their morphology and vegetation cover tend to reduce soil erosion while their metabolism influences soil microbial activity and chemical properties (Bochet et al. 2006; Cerda et al. 2021; Fromin et al. 2020; Gimeno-Garcia et al. 2001). Spatial distribution of plant species is typically more affected by abiotic factors at larger scales while biotic interactions dominate at smaller scales. However, in semi-arid Mediterranean scrubland, Pescador et al. (2020) reported that the higher soil heterogeneity may in fact play a greater role than biotic factors at smaller scales.

The endemic shrub *Anthyllis hermanniae* subsp. *Melitensis* (Yellow Kidney Vetch) is commonly found in the Maltese garrigue associations described above, and it distinguishes them from similar non-endemic ones (Brullo et al. 2020). Its optimal conditions are in areas exposed to wind and sea spray, but it is also commonly found inland, although with smaller morphologies

(Brullo et al. 2020). It is a typically dense and short shrub (<50cm) displaying twisted branches with spiny ends, dark green, sericeous 1 to 3-foliolate leaves, and yellow flowers appearing between April and June (Brullo & de Carlo 2006). Several other shrub species are found in the Maltese garrigue and characterize the vegetation assemblages (see section 2.1.3.1).

2.1.3.3 Grazing

The Intermediate Disturbance Hypothesis (IDH) postulates that the relationship between species diversity and ecosystem disturbances is hump-shaped, with species diversity being maximized at an intermediate level of disturbance (Gao & Carmel 2020a). However, there has been discrepancies between its predictions and empirical results, notably in dry climates. The Milchunas–Sala–Lauenroth (MSL) model (Milchunas, Sala & Lauenroth 1988) generalized the relationship between grazing intensity, moisture level, and the evolutionary history of grazing in grasslands (Figure 2.2). It predicts that in dry areas, diversity is always negatively correlated to grazing intensity, but that the IDH holds in wet areas, while a long coevolution with grazing attenuates the effects of grazing on diversity everywhere, though with a minimal effect in semi-arid areas with a long history of grazing (Gao & Carmel 2020a; Milchunas, Sala & Lauenroth 1988).

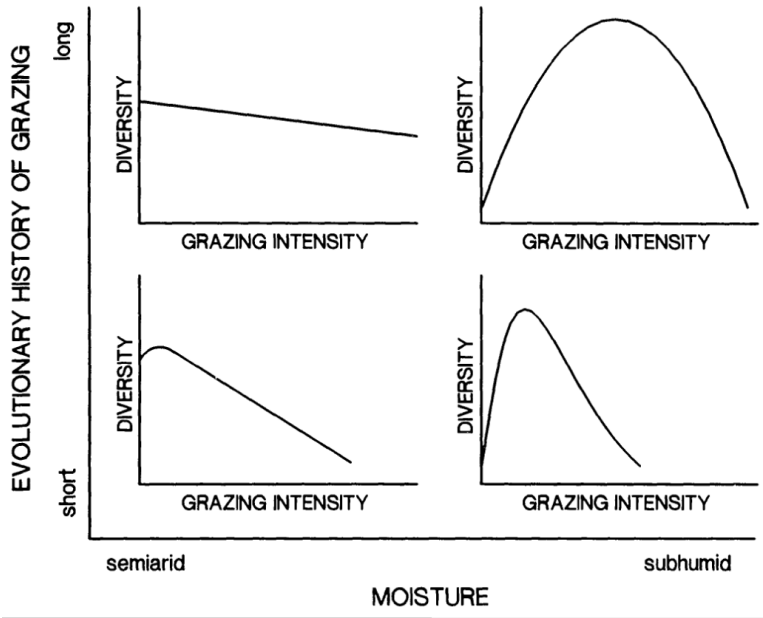


Figure 2.2. The Milchunas-Sala-Lauenroth model, describing the relationship between plant diversity and grazing intensity across gradients of moisture and evolutionary history of grazing. From Milchunas, Sala & Lauenroth (1988, p.95).

However, the MSL model assumes that a specific plant community will have a single level of diversity at each level of grazing intensity but does not account for the possibility that there may be multiple possible stable community compositions at a given grazing intensity, which can be influenced by past grazing practices or environmental events (Cingolani, Noy-Meir & Diaz 2005). To overcome these limitations, the latter authors proposed three main improvements to the MSL model: (i) the moisture gradient was generalized to a productivity gradient; (ii) an alternative mechanism was proposed to explain the weak diversity response in long-grazed, low-productivity communities, besides convergent selection by grazing and drought; (iii) situations with discontinuous, irreversible, and divergent responses to grazing intensity were allowed. It highlighted that in general, communities with a long history of grazing rarely witness such thresholds because of a higher proportion of "grazing-plastic" and "grazing-dependent" groups allowing to reverse changes made by grazing (Alrababah et al. 2007, p.262).

On the other hand, communities with a short history of grazing are more likely to experience irreversible transitions to alternative ecosystem states, with multiple factors influencing the nature of the transition (Alrababah et al. 2007, p.262).

Gao and Carmel (2020b) performed a meta-analysis (n=259) at the global level to review the effects of grazing on plant richness. In general, grazing significantly increased plant richness in wet areas, whereas it influenced species abundance and composition rather than species richness in dry areas. However, the importance of grazing on species richness changed with spatial scale; grazing variables had a stronger influence at the site level because of local homogeneity, while aridity and ecosystem variables became more important at continental and global levels.

In the Mediterranean Basin, Sternberg et al. (2020) reported that annual precipitation was the main influence of vegetation cover for most plant functional groups in a semi-arid environment, while the effects of grazing were significant only during relatively wetter years, in accordance with the expanded MSL model. Alrababah et al. (2007) did not find any effect of grazing on grassland plant species richness in the eastern Mediterranean Basin, but on species composition and abundance instead, validating the observation reported by Gao and Carmel (2020b). However, they found that increased grazing reduced vegetation cover which in turn caused a positive feedback loop with increased soil erosion.

Sternberg et al. (2020) also examined variations in the effects of grazing on plant types according to their life cycle and morphology in Mediterranean environments. Tall annual grasses are particularly affected by grazing which decreases their cover, but shorter, prostrate and rosette-forming species experience a compensatory increase instead. In contrast, some hemicryptophytic species and geophytes (e.g. *Asphodelus ramosus*) were found to be less

affected by grazing because of adaptations such as buried perennating buds reducing the dependence on seed production, physical or chemical defences, and secondary chemical compounds.

The Mediterranean landscapes have a long history of anthropogenic influences, among which livestock grazing is significant since at least 5000 years (Sternberg et al. 2020). In Malta, Schembri and Lanfranco (1993) suggested that grazing by domestic goats and sheep might have played a significant role in shaping the landscape. Unused or in excess land may have been designated for grazing which eventually turned into overgrazing, thereby suppressing the growth of maquis and woodland vegetation. In fact, local pollen history indicates that the combination of grazing and fire, used as a land-clearing tool during the early Holocene, has contributed to open the landscape, gradually making garigue and steppe the dominant vegetation types on the Maltese islands (Farrell, Hunt & McClung 2020). Recently, as it became uneconomical to maintain large herds of goats and sheep in Malta, farmers were encouraged to invest in enclosed cattle, furthering the decrease in goats and sheep populations and their related open grazing, thereby causing a slow process of shrub regeneration in the southern and south-eastern parts of Malta (Attard 2008).

2.2 The wild European rabbit

2.2.1 Taxonomy and biology

The European rabbit (*Oryctolagus cuniculus*) is a small mammal belonging to the family Leporidae and the order Lagomorph. Fossil evidence suggests that the genus *Oryctolagus* evolved on the Iberian Peninsula approximately 3.5 million years ago, while the oldest known fossil of the species *O. cuniculus* has been found in the same region and is estimated to be around 0.6 million years old (Lopez-Martinez 2008). The European rabbit is the only species

in its genus to have survived the Late Pleistocene period, which is thought to be due to its ability to adapt to changing climates. The modern subspecies of the European rabbit, spp. *cuniculus* and spp. *algirus*, are thought to have emerged on the Iberian Peninsula during the last glacial maximum (Ferrand & Branco 2007). The European rabbit has been widely dispersed throughout the world due to its high reproductive rate, its phenotypic plasticity, and usefulness to humans as a source of food, fur, and bones (Flux & Fullagar 1992; Marin-Garcia & Llobat 2021).

Oryctolagus cuniculus is a generalist herbivore in Mediterranean ecosystems that feeds mainly on grasses and herbaceous plants (Delibes-Mateos 2020). It uses adaptive feeding strategies to maximize the quantity and quality of nutrients it obtains, such as eating barks during dry periods and occasionally consuming mildly toxic plant species (Marin-Garcia & Llobat 2021). European rabbits have high nutritional requirements that may not be met in all ecosystems, but their plasticity allows them to compensate by adapting their feeding strategies seasonally. Marin-Garcia and Llobat (2021) mention that nutrition, particularly protein, can limit its abundance because poor nutrition can hinder rabbit survival through increased predation risk and susceptibility to epizootic events. Due to its high metabolism, food shortages have been observed to reduce heart rate in rabbits (Eisermann et al. 1993) and affect seasonal population variations (Kontsiotis, Bakaloudis & Tsiompanoudis 2013).

Leporids also practice caecotrophy, which involves a digestive system with a developed cecum that separates food into finer nutritive particles (soft faeces) and larger, poorly digestible fibrous material (hard faeces). Hirakawa (2021) provides a detailed description of the caecotrophy process in *Oryctolagus cuniculus*. Fresh food is typically consumed during the evening, filling the digestive tract and activating the separation mechanism. In the morning, the resulting hard faeces are excreted and reingested in parallel to the onset of soft faeces

production. The latter are subjected to intestine microbial action and coated with a protective mucous membrane, before being excreted and directly reingested during the middle of the day. The mucous membrane on soft faeces allows for continued microbial activity during the second digestion, optimizing carbohydrate breakdown. The stomach content decreases during the afternoon until all ingested material has cycled at least once, and the cycle starts again, with leftover hard faeces discarded during night feeding. Soft faeces are rarely found outside the animal, while hard faeces are commonly found in the open. Mancilla-Leyton, Gonzales-Redondo and Vicente (2013) noted that caecotrophy also influences seed germination and dispersal by the European rabbit. Hard faeces have been found to be coated as well, but with odoriferous anal-gland secretions used for information exchange (Sneddon 1991). They are either deposited at random during feeding, or as dunghills called "latrines" mainly used in social interactions.

2.2.2 Ecology of the European rabbit

Latrines are centres of exchange for territorial, reproductive, and hierarchical information between and inside groups of rabbits, while their use has been found to increase with dominance status (Sneddon 1991). The European rabbit is indeed a territorial species with a hierarchical social structure, forming groups made of a dominant buck, several females and subaltern males living in underground burrows called warrens, which range from the simple hole to complex networks of chambers (Cowan 1987). However, the size and composition of a group also depend on how closely nest sites are located together and how limited the availability of nest sites is. The home range of a warren vary between 0.5ha and 4ha according to the landscape, predation pressure, refuge and food availability, and weather harshness including flood susceptibility (Lombardi et al. 2007; Moseby et al. 2005). Warrens are not exclusive, and rabbits use multiple warrens of a territory over time (approx. 2.5 warrens/rabbit;

Moseby et al. 2005). Factors affecting warren location selection are complex and depends on landscape- and patch-level fragmentation, soil types and depth, resources availability, predation pressure, vegetation cover and root structure (Dellafiore et al. 2008; Palomares 2003). Kontsiosis et al. (2013) found that soil hardness, measured as the hardness of the soil cover, was the main predicting factor of European rabbit density on Lemnos Island, Greece.

At the landscape level, Dellafiore et al. (2008) observed in Spain that "the probability of finding a warren increases as the perimeter-area ratio [of patches] decreases" (p. 377), implying that European rabbits tend to prefer a less fragmented landscape with few larger patches to an area with many smaller patches, as well as less edge effects. However, Monzon et al. (2004) found in Portugal that European rabbits favoured landscapes with higher vegetation fragmentation and higher edge effect, which increased resources diversity in the home range. In the same study, European rabbits were found to be more abundant in scrubland than in forest. Similarly, Dellafiore et al. (2008) observed in Spain that warrens were preferably dug in areas with higher shrub understorey, which provides better shelter from predators, reduces soil erosion through increased cover and supports soil stability through its root system (Dellafiore et al. 2008). The differences in habitat selection caused by landscape fragmentation and edge effects have been considered more related to predation pressure and availability of shelter around warren patches than to food availability (Dellafiore et al. 2008).

Schweizer et al. (2016) conducted a meta-analysis of the effects of rabbit eradication over vegetation in islands. They found that plant richness and vegetation cover often increase after eradication, with the increase intensity being proportional to the grazing intensity pre-eradication. Similar outcomes were observed for native and exotic plant species, even though the latter may experience a population explosion following the release from intense grazing pressure.

Fernandez (2005) modelled wild rabbit abundance as a function of habitat descriptors at three nested spatial scales in the Iberian Peninsula. The study found that rabbit abundance was linearly associated with shrublands-pasture ecotones and non-linearly associated with shrub cover at the home-range scale (3ha). This was explained by the balance required by the rabbit between areas of refuge and food availability in its home-range. At the landscape level, a positive effect of distance to freshwater (streams, ponds and lagoons) on rabbit abundance was observed, which was hypothesized to be related to higher groundwater table and more year-round growth of the herbaceous layer.

Carvalho and Gomes (2004) found a hump-shaped relationship between relative rabbit density and herbaceous cover in north-west Portugal, used here as a food availability index. They also found a linear positive correlation with tall scrub cover where it was above 20%, but it was absent and replaced instead by a linear positive correlation with gaps in rocks where the tall scrub cover was below 20%. They interpreted the effect of small herbaceous cover values on rabbit density as a food limiting factor, while high herbaceous cover values were interpreted as a sign of open landscape and therefore, as a refuge limiting factor.

Planillo and Malo (2018) highlighted that the effects of human infrastructure and its characteristics had a larger influence on rabbit density than environmental parameters at the range scale in central Spain. In particular, the width of road verges in association with the traffic flow were found to be the main factors influencing rabbit density. Roads with low traffic and large vegetated verges sustained higher rabbit densities in areas with a natural environment that would not have supported these densities otherwise. A study of the home-range of European rabbit in suburban and urban areas of Germany (Ziege et al. 2020) found that the local structural heterogeneity had a larger influence on the home range of European rabbits than habitat fragmentation. Radio-tracked rabbits were observed to move underground below

roads and other traditional barriers in the city, but they were restraining their activity to the minimum area (0.5ha) offering both food sources and shelter instead of those maximizing food availability.

Beja, Pais and Palma (2007) studied the impacts of shrub vegetation structure and composition on European rabbit latrine density in south-west Portugal. They discovered that the ground-level shrub vegetation cover density had a significant effect on latrine density. The latter was positively associated with the proportion of shrub species displaying sparse ground-level vegetation but dense canopy (*Cistus ladanifer),* as well as with important herbaceous cover within shrub patches. Conversely, they identified a negative correlation between latrine density and the vegetation cover of shrub species with dense ground-level vegetation (ericoid and broad-leaved species).

In the Canary Islands, Cabrera-Rodriguez (2006) looked at the influence of the number of shrubs and the proportion of various land covers including herbs on relative rabbit density measured with dung pellet counts. The author found a positive correlation between herb cover and rabbit pellets, while there was a lower number of pellets at higher altitudes. There were more rabbits in coastal and transformed areas, which had a higher grass cover, than in forests where grass was scarcer. In the same context, Cubas et al. (2021) reported that the strongest determining factor of rabbit density across diverse habitats was vegetation structure, described here by tree cover which had a negative relationship to rabbit density. Rock cover and slope were also found to have a negatively relationship to rabbit density, albeit to a lesser extent and depending on the type of habitat.

Calvete et al. (2004) studied the relationships between rabbit relative density (using latrines and pellets count), environmental and vegetation variables in an agricultural area of north-east

Spain using generalized linear models. They found that the distribution of rabbits was negatively related to strong topography and hard soils, while it was positively related to shrub cover and shorter distances between refuge vegetation and alternated cereal crops. They also found a high variability of rabbit relative density between sites deemed suitable to higher rabbit density by their model, which points to unknown factors limiting reproduction or re-colonisation such as diseases or ecological connectivity.

2.2.3 Abundance measurements

Abundance can be measured through absolute or relative density estimations. Absolute density refers to the number of individuals within a given area or volume, while relative density is derived from a biological index that bears correlation with absolute density and is used for comparison of populations (McComb et al. 2010). Absolute density indicators often consist in the direct observation of individuals by unit of area, length, or volume, whereas relative density indicators are made of indirect observations of individuals, typically through abundance indices such as faecal pellets, nests, or other reliable indicators of space occupancy (Elzinga 2001). Hopkins and Kennedy (2004) suggest that relative abundance indices can be used to discern broad trends in small mammal populations, provided they are calibrated against absolute abundance indices. According to Palomares (2001, p. 584), an "uncontrolled counting of pellets [relative to a direct method] is not sufficient to accurately estimate [absolute] rabbit abundance, […] yet […], it may provide a satisfactory approximation of relative rabbit abundance".

In a review conducted by Fernandez-de-Simon et al. (2011), significant correlations were found between various direct and indirect methodologies used to estimate the abundance of *Oryctolagus cuniculus* in Spain, such as night counts from driven transects and pellet-count indices. They recommend counting standing crop pellet (STP) as one of the most trustworthy indirect methods for gauging rabbit abundance in Mediterranean ecosystems during the

summer season. This methodology involves counting of dung pellets in quadrats distributed across an area following a pre-determined sampling design, yielding a pellet density index (pellets/area). Putman (1984) extensively discusses the diverse applications of faecal pellets in mammalian ecology, emphasizing the significance of considering decay rates when using STP counts for estimating absolute abundance. European rabbit dung pellets can endure in situ for several months, largely depending on complex interactions between climatic conditions, vegetation and local fauna such as insects (Cubas et al. 2021), with Wood (1988) reporting a 95% survival rate of pellets after 14 weeks in spring over five years. During hotter seasons, desiccated pellets become lighter and thus more easily displaced or damaged by rain impact, while slope and vegetation cover elevate the chances and distance of displacement by wind and surface run-off. As such, STP count should primarily be used for intra-seasonal pellets density comparison, unless differential seasonal decay rates are considered (Iborra & Lumaret 1997). Cubas et al. (2021) confirmed that the variations in decay rates do not hinder STP counts from being comparable with cleared-plot pellet counts across locations.

The issue of spatial heterogeneity caused by latrines is also explored by Putman (1984), wherein he suggests that pellets deposited in latrines have stronger scent marks because they serve to signal reproductive, territorial, and dominance status between groups and individuals, while non-latrine pellets present lower scent gradients as they are not used for communication. Pellets in latrines incite further deposition, while non-latrines pellets do not affect it significantly (Sneddon 1991); consequently, STP counts which include latrines pellets are skewed due to the clustered spatial distribution of pellets in latrines. To counter this, Mutze et al. (2014, p. 242) developed a method to adjust the average pellet density outside of latrines “for the expected number and size of latrines in that area", which was experimented in Malta (Tranchant 2022). This later study enabled the use of faecal pellet counts to estimate absolute

rabbit abundance; however, a few biological parameters were missing for *Oryctolagus cuniculus* in this local context, which reduced the accuracy of the absolute estimation. Nonetheless, non-latrines faecal pellet counts proved to be a resource-efficient method to estimate relative rabbit abundance and compare it across sites.

2.2.4 The European rabbit in Malta

The European rabbit is believed to have been dispersed to the Maltese archipelago by Phoenicians seafarers during Antiquity (Cassar 1994). The earliest record of the species in Malta is dated between 100 B.C. and 100 A.D. from the Juno temple in Tal Silg (Masseti & De Marinis 2008). There is a paucity of ecological research on the wild *Oryctolagus cuniculus* population of the Maltese archipelago, and their abundance and distribution are still largely undocumented. Wild populations are currently managed through a cultural and political framework with the Protection of Wild Rabbit Regulation, subsidiary legislation 549.90 of 2014. There are around 3000 hunters licensed for rabbit hunting, which is allowed between the 1st June and 31st December, however, there are conflicts around the potential exploitation of the rabbit season to hunt birds instead (Birdlife 2019).

Beside humans, there are few potential predators of *Oryctolagus cuniculus* in the Maltese archipelago (Table 2.2). Among these, predation by *Mustela nivalis* could be consequential, as could be the seasonal airborne predation from migratory raptors like *Circus aeruginosus*. However, none of these ecological interactions have been scientifically assessed. On the islands of Malta and Gozo, the presence of weasels has been substantiated but not quantified (Cauchi 2014; Lanfranco 1969), while it has not been confirmed on Comino.

Table 2.2 *Potential predators of Oryctolagus cuniculus in Malta.*

Maltese name	English name	Scientific name	Environment	Location	Rabbit preys
Ballottra	Weasel	*Mustela nivalis*	Terrestrial	Malta, Gozo, Comino	Juveniles, sicks
Nemes/farrett	Ferret	*Putorius putorius furo*	Terrestrial	Malta, Gozo	All
/	Green Whip Snake	*Hierophis viridiflavus*	Terrestrial	Malta, Gozo, Comino	Juveniles
Lifgħa	Leopard Snake	*Zamenis situla*	Terrestrial	Malta, Gozo	Juveniles
/	Western marsh harrier	*Circus aeruginosus*	Aerial	Malta, Gozo, Comino	Juveniles, sicks
/	Black Kite	*Milvus migrans*	Aerial	Malta, Gozo, Comino	Juveniles, sicks, carrion
/	Red-footed Falcon	*Falco vespertinus*	Aerial	Malta, Gozo, Comino	Juveniles, sicks

Apart from predation, disease also acts as a significant natural regulating factor. The Rabbit Haemorrhagic Disease (RHD) is due to a contagious calicivirus and propagated through interactions between infected media and uninfected individuals, either directly or through insects (Abrantes et al. 2012). The *Myxoma* virus, which instigates myxomatosis, is a pox arbovirus, typically non-lethal to lagomorphs within its indigenous American range but with mild to lethal effects on European populations. These two diseases have been harnessed by humans for population control, particularly in Europe and Australia, where they have kept populations at approximately 15% of their initial size (Agriculture Victoria 2021). These diseases provoked a substantial decrease in the domestic and wild rabbit population in the Maltese archipelago three decades ago, and they are managed today via national vaccination initiatives for domestic animals and livestock (Veterinary Regulation Directorate 2021). However, recent virus mutations have led to an upsurge in fatalities within the Maltese wild populations of *Oryctolagus cuniculus*, which are not safeguarded by vaccination (Muscat 2021).

2.3 Unmanned Aerial Systems & Land Cover

2.3.1 General definitions

Unmanned Systems (US) can be defined as:

> powered vehicles (air, maritime, and ground) that do not carry a human operator, can be operated autonomously or remotely, can be expendable or recoverable, and can carry a variety of payloads depending on their type, functionality, operational characteristics, and mission objectives. (Valavanis & Vachtsevanos, 2015, p. 11)

According to the ICAO (2011), US are typically made of:

- An Unmanned Vehicle (UV) or drone, the primary component of the US which is defined as a powered reusable vehicle without an on-board human pilot. The term "reusable" distinguishes UVs from the various guided weapon systems.
- A Ground Control Station (GCS), the hardware/software device used by humans to monitor and control the drone remotely (and not only from ground).
- A Command and Control (C2) link, the hardware ensuring the transmission of data between the UV and the GCS.

Furthermore, the term unmanned system generally refers to the presence of a human in the system, albeit only for programming an autopilot (Dalamagkidis 2015). US can operate in various environments, such as air, ground, and water, and those designed for flight are referred to as Unmanned Aerial Systems (UAS), which utilize Unmanned Aerial Vehicles (UAV). In the context of remote sensing (RS) applications, UAS typically incorporate navigation, orientation, and sensing equipment also known as 'payloads' (Mejias et al. 2015). The Navigation System (NS) employs real-time position, velocity, and altitude parameters to inform the Flight Control System (FCS) of the aircraft's autopilot at a high frequency and moderate accuracy.

Conversely, the orientation system records these parameters with higher accuracy but lower frequency, because they are usually stored with the captured data to provide information on the sensor's position and orientation for post-processing (Colomina & Molina 2014). Today, due to miniaturization, the NS and orientation system are often integrated into a Hybrid Navigation System (HNS). The sensing payload comprises one or more lightweight electro-optical or radio-wave sensors from a broad range of available options. The most common sensors include visible, multispectral, and hyperspectral cameras, thermal imaging systems, laser scanners, and radars (Colomina & Molina 2014). A sensor's properties (such as physical dimensions, focal length, and image width and height in pixels) combined with its altitude determine its image resolution, referred to as Ground-Sampling Distance (GSD). GSD is expressed as the 'on-the-ground distance per image pixel', with contemporary GSDs achieving ultra-high resolutions of less than a centimetre per pixel (Yao et al. 2019).

2.3.2 History of UAS

UAV have been around since Antiquity, with the use of various technologies to automate the flight of an object (Dalamagkidis 2015). However, the modern concept of UAS as drones was developed for military applications at the beginning of the 20th century following new communication and information technologies. It reached the research community in the 1970s and became eventually popularized in all strata of society from the beginning of the 21st century (Colomina & Molina 2014). Today, they are the subject of a large and varied body of national and international regulations aimed at better defining and facilitating their use in relation to pre-existing airspace usage (EASA 2009; ICAO 2011).

In the last decade, UAS applications for remote-sensing have surged in environmental sciences by enabling the acquisition of ultra-high resolution datasets with commercial-grade equipment, effectively filling a gap between small-scale satellite imagery and large-scale fieldwork data

(Anderson & Gaston 2013; Nowak, Dziób & Bogawski 2019). UAS also provide greater flexibility in terms of temporal resolution as they allow to decide when to acquire data, unlike previous remote-sensing technologies that have periodicity constraints (e.g. satellites). While various specialized payloads have been adapted to UAS from other remote-sensing technologies (e.g. LiDAR), the standard RGB camera of most commercial UAV still allow to derive spectral, textural and geometrical information from a collection of photographs through a process known as 'Structure from Motion' (SfM) (de Castro et al. 2021).

2.3.3 Structure from Motion

The combination of UAS flexibility and SfM performance in modern software has led to a wide adoption of UAS-SfM products in environmental sciences, with applications ranging from inventories to geological mapping (Anderson & Gaston 2013; Iglhaut et al. 2019; Linchant et al. 2015; Nowak et al. 2019). Iglhaut et al. (2019) provide a thorough description of the theory behind SfM. Stereophotogrammetry has its roots in the concept of human binocular vision, where depth can be derived from two points with known relative positions. However, depth, volume, or 3D features can also be deduced from a single observing point if either the observer or the object is in motion. SfM combines both principles of binocular vision and changing perspectives of a moving object by overlapping 2D images to generate orthomosaics or 3D models. Images can be acquired through any means as long as they provide total coverage of the desired surfaces with sufficient front and side overlaps between them, depending on the objective. Iglhaut et al. (2019) specified that SfM distinguishes itself from traditional photogrammetry by three key aspects:

(i) features in images can be automatically identified and matched even when they have different scales, viewing angles, and orientations, which is particularly useful for small and unstable platforms;

(ii) camera position or ground control point information is not mandatory;

(iii) camera calibration can be automated. In reality, SfM is only one step of a lengthier process called 'Structure-from-Motion with Multi-View Stereo photogrammetry' (SfM-MVS).

The process starts with matching images using a Scale-Invariant Feature Transform (SIFT) algorithm or similar, which identifies distinctive keypoints (pixels or sets of pixels) in images and associates them with numerical descriptors insensitive to changes in scale and orientation. The second step of the process is the actual SfM, where the position and orientation of the camera for each image is calculated simultaneously with the generation of a Sparse Point Cloud (SPC), comprised of all keypoints matched in all images. These first two steps are often performed together in most commercial software. At this stage, the SPC is scaled to determine the relative distances between keypoints, using either calculated camera positions, which are less accurate as these positions are only estimated, or using Ground Control Points (GCPs), which are visual markers placed on the ground prior to photography and whose coordinates have been accurately recorded using specialized equipment. Thirdly, Multi-View Stereo (MVS) algorithms are used to densify the SPC by clustering the images based on their position to generate more keypoints, resulting in a Dense Point Cloud (DPC) which is generally the main output of the SfM-MVS process. The resulting points of the SPC and DPC are made of spatial coordinates (x, y, z) and spectral information, such as RGB and HSV values in the case of visible range imagery. The DPC can then be cleaned and further processed into Digital Elevation Models (DEM), orthomosaics, or 3D products such as textured meshes.

SfM-MVS products are affected by two types of accuracy: geometric and positional. Geometric accuracy refers to errors in distances within the product, while positional accuracy refers to errors in the location of the product within a defined Coordinate Reference System (CRS). In

the context of UAS studies, the accuracy of SfM products depends on a large number of parameters that are generally classified according to the stage at which they can be adjusted (Iglhaut et al. 2019). These stages can be simplified into three categories: pre-flight parameters, environmental conditions, and data processing parameters. These parameters must also be adjusted according to the objectives of the survey, but general guidelines apply to obtain a quality 3D product.

Pre-flight parameters refer to all adjustments made to the UAV, its camera, and the flight plan before image capture. These include flight altitude and speed, the "exposure triangle," which refers to the ISO, shutter speed, and aperture of the camera, the use of GCPs and their placement, the vertical angle and focal length of the camera, and the degree of overlap between images. For applications where geometry is important (horizontal and vertical distances, surfaces, volumes), the literature cited at the beginning of this section recommends an overlap of at least 80% in all directions, as well as an oblique camera angle (70°) and a flight speed adjusted according to the altitude and the photography interval. Similarly, the exposure parameters of the camera should be known before setting the flight parameters, as the aperture and ISO should be proportional to the flight speed and altitude to reduce motion blur. Cunliffe et al. (2016) emphasized the importance of using non-nadir (oblique) imagery to achieve fine resolution with SfM derived from UAV images, as it enhances the number and precision of ground points, leading to higher quality DTMs. Howell et al. (2020) supported these findings while suggesting that research objectives and resource limitations should be the primary determinants of flight parameters for shrub measurements since various flight parameters did not significantly affect their model accuracy or the correlation between modelled and measured shrub heights. Compressing images (e.g. jpg) at the time of capture by algorithms

native to each UAV reduces the positional accuracy of products by about 2 cm compared to uncompressed images (Kalacska et al. 2020).

The use of GCPs or integrated positioning modules is recommended, as this significantly increases the geometric and positional accuracy of products (Kalacska et al. 2020; Liu et al. 2022; Sanz-Ablanedo et al. 2018). GCPs should be placed homogeneously to form triangles in the area of interest with GCPs along the edges as well as in the centre. Although the required density of GCPs depends on the resolution and accuracy objectives as well as on the scale of a project, approximately 1 to 3 GCPs per surveyed hectare has been deemed sufficient to achieve stable 3D accuracy (Ulvi 2021; Yu et al. 2020). GCPs are typically collected using topographic Global Navigation Satellite System (GNSS) receivers (e.g., total stations), but accurate positions of the camera can also be used as GCPs if the UAV is equipped with an on-board Real-Time Kinematics (RTK) module connected to a local Continuously Operating Reference Station (CORS) network (Kalacska et al. 2020).

Environmental conditions such as illumination (i.e., sun position, cloud cover), wind, topography, and terrain feature characteristics (texture, repetitive patterns) all affect image quality as well as flight plan choices. Illumination conditions should be as constant as possible between images in the same dataset, and flights should be conducted with as little wind as possible to improve the ability to identify keypoints across images by reducing motion and spectral variations in the photographed objects. In the case of RGB imagery used for structural analysis, an even illumination with full cloud cover is recommended as it minimizes shadows, while for multispectral imagery, clear skies are preferred as it minimizes the atmospheric influence on the spectral bands and maximizes information in the infrared spectrum (Iglhaut et al. 2019).

Finally, data processing parameters are set within the SfM-MVS software. These are generally specific to each software package and use proprietary algorithms that function as 'black-boxes' but are nevertheless flexible. Tinkham and Swayze (2021) have shown that the *Quality* and *Depth Map Filtering* (DMF) parameters for the generation of a DPC in Agisoft Metashape software had significant effects on positional and vertical geometric accuracy. Both *High* and *Ultrahigh Quality* settings produced similar point cloud density with satisfactory accuracy and precision but showed a significant difference in processing performance. Any setting above *Mild* for DMF resulted in significant degradation of vertical information and reduction in point cloud density, whereas there was virtually no difference between *Mild* and *Deactivated*.

2.3.4 DTM from UAV imagery

The main products derived from the DPC are the different DEMs and the orthomosaic. The DEM derived directly from the DPC is called a Digital Surface Model (DSM), which represents all the highest points of the DPC. It is also possible to create a Digital Terrain Model (DTM) that represents only the topography of the bare terrain without any features present on the surface such as vegetation or buildings. Subtracting the DTM from the DSM allows the creation of an Above-Ground Model (AGM), which represent the relative elevation of all features above the ground. These models are essential to derive vertical information from UAV vegetation surveys.

Generating DTMs from UAV photogrammetric point clouds however can be challenging because the passive optical sensors cannot penetrate aboveground objects, such as vegetation cover, preventing or severely limiting the presence of ground points below them (Salach et al. 2018; Stroner et al. 2021). The main approach consists in filtering points above the ground based on their geometry, colour, texture, or combinations thereof (Anders et al. 2019; Cunliffe, Brazier & Anderson 2016; Jensen & Matthews 2016; Jiménez-Jiménez et al. 2021; Howland et

al. 2022; Serifoglu Yilmaz & Gungor 2018; Zeybek & Şanlioglu 2019). The removal of non-ground points leaves gaps in the DPC that are eventually filled by interpolation to produce a continuous surface that forms the DTM. The degree of errors in the DTM affects the vertical accuracy of the resulting AGM, which can ultimately prevent its use. A common method to assess the accuracy of the DTM or AGM is by using measurements made in the field (ground-truthing) eventually compared to digital measurements (Salach et al. 2018).

An alternative solution is to obtain a DTM generated by LiDAR acquisition, which provides a point cloud where ground points are already classified as such because it is able to detect partial penetrations of the laser (Stroner et al. 2021). However, there can be significant differences in resolution between LiDAR and UAV-photogrammetric datasets. In Malta, terrestrial LiDAR scans are made publicly available by the Planning Authority (PA) and cover the entire archipelago at a resolution of 15 cm (MEPA 2013).

2.3.5 UAV for vegetation surveys

Recent advancements in UAV technology have significantly improved the accuracy and efficiency of vegetation mapping, particularly in scrub environments. Vafidis (2021) demonstrated that using UAVs with a GSD of approximately 2.5 cm/pixel enabled the classification of scrubs based on height after generating an Above-Ground Model (AGM). This allowed for the calculation of the surface area occupied by each height category. Cunliffe et al. (2016) used arbitrary relative elevation thresholds from the AGM layer to distinguish bare ground surfaces from vegetation (>0.015m), as well as to better separate herbaceous plants from woody ones (>0.2m). To further enhance the accuracy of vegetation classification, Prosek and Simova (2019) found that fusing multispectral imagery and SfM vertical data from a single sensor yielded better results at the species level compared to using multispectral imagery alone. Li et al. (2021) added that the combination of spectral and textural indices derived from

RGB imagery facilitated image segmentation and classification, particularly when dealing with high spectral similarity between shrublands and surrounding areas. Notably, the Excess Green Index and RGBVI were effective in distinguishing vegetation from non-vegetation surfaces.

When examining gaps in vegetation, Charton et al. (2021) highlighted the effectiveness of UAV models for accurately mapping and measuring gaps, enabling comparisons of gap frequency and size with other landscape variables. Addressing the issue of larger pixel sizes in aerial imagery, Hu et al. (2018) noted that the "mixed-pixel effect" could reduce the efficiency of binary classification. However, employing object-based classification can mitigate this issue by grouping similar neighbouring pixels into homogeneous objects.

2.3.6 Image classification

In remote sensing, acquired imagery is often processed to detect meaningful objects or categories, eventually used for further analysis. Image classification refers to the process of attributing classes or categorical information to subsets depicted in the image, such as land cover classes (Richards 2022). Unsupervised classification implies the automatic definition of classes by an algorithm through the identification of distinct pixel vectors, and it is typically used when there is no prior knowledge of the type and presence of land covers. It is opposed to supervised classification which entails the preliminary definition of classes by the user through the provision of training samples (Richards 2022). These training samples allow to determine “boundaries between information classes in the spectral domain” (Richards 2022, p. 98). Other sources of information can be supplemented to aid in the definition of those classes, such as elevation data.

Another common dichotomy is the use of either pixel-based or object-based image classification approaches. The former classifies pixels of the image by assigning them

individually to a class, while the latter requires a previous segmentation of the image into meaningful objects (i.e., groups of pixels) based on textural, spectral or topographical similarities, before classifying these objects into classes (Sibaruddin et al. 2018). Ultra-high resolution imagery obtained by UAS is often processed through object-based image classification because individual pixels rarely represent meaningful information at this resolution and neighbouring pixels are not considered during pixel-based classification, whereas instead, object segmentation allows to group similar neighbour pixels according to the scale of features to be detected (Hossain & Chen 2019; Sibarruddin et al. 2018). Various approaches to segmentation exist, including multiscale segmentation which consider multiple scales of objects in a single image, as it is often the case in ultra-high resolution imagery (Hossain & Chen 2019). Object-based classification algorithms are often reported to show improved accuracy over pixel-based methods; however, resource requirements may vary significantly, and research objectives can ultimately determine the most appropriate approach (Behane et al. 2017; De Giglio et al. 2019; Lu & Weng 2007; Whiteside et al. 2011)

A large range of classifier algorithms is available, either as parametric when they assume normally distributed datasets (such as Maximum Likelihood), or non-parametric when the statistical distribution of the data is unknown or not normal. Among the latter, machine-learning algorithms have become prominent in recent years due to their capacity to "model complex class signatures, accept a variety of input predictor data, and not make assumptions about the data distribution" (Maxwell, Warner & Fang 2018, p.2784). Among these, single or ensemble decision trees (e.g. Random Forest, RF) and Support Vector Machines (SVM) algorithms are commonly used in remote sensing classification and are typically available in commercial software packages (Basheer et al. 2022). Random Forest (RF) classifiers use a large number of single decision-trees as ensemble, randomly subset unknown objects, and

submit their class assignation to random separate decision trees, eventually comparing and selecting the best one, with the advantage of providing a measure of each band's contribution to the classification (Belgiu & Drăguț 2016). On the other hand, SVM classifiers use kernel functions to fit a hyperplane that splits the dataset into a predetermined number of classes through iterative optimization of a decision boundary (Mountrakis, Im & Ogole 2011). They are particularly adapted to an imbalanced number of training samples per class and offer a good balance between overfitting and generalization, however, they hardly provide any measure of the bands' importance in the classification, and require complex parametrization, in particular with the selection of the kernel function (Maulik & Chakraborty 2017). For the classification of UAV imagery for vegetation and land cover surveys, SVM classifiers have been reported to outperform RF both with pixel and object-based analysis (Basheer et al. 2022).

Post-classification processing can often improve its accuracy or its relevance through noise filtering and integration of secondary data into a reclassification process (Lu & Weng 2007). Classification accuracy assessment is the most common used method to assess and compare the accuracy of different classifications or post-processing methods (Lu & Weng 2007). It involves the creation of a reference dataset containing the type of objects to be classified, and ideally selected through stratified random sampling (Richards 2020). The reference dataset is then classified manually through ground truthing or remote visual inspecting by an expert. It is eventually compared to the output of the classification procedure to derive an error matrix, which provides a number of key statistics for the analysis of accuracy (Congalton 1991):

I. Commission error & user's accuracy

Commission error refers to Type I error (false positive). It is computed from the classified data and represents the number of objects incorrectly included in a given

class. The user's accuracy represents the probability that an object classified on the image correspond to the same class in reality.

$$\textbf{Commission error} = \frac{Total\ number\ of\ incorrect\ objects\ in\ a\ class}{Total\ number\ of\ objects\ in\ a\ class}$$

$$\textbf{User's accuracy} = 1 - commission\ error$$

II. Omission error & producer's accuracy

Omission error refers to Type II error (false negative). It is computed from the reference data and represents the number of reference objects omitted in a given class. The producer's accuracy represents the probability that an object in reality will be correctly classified as such in the image.

$$\textbf{Omission error} = \frac{Total\ number\ of\ incorrect\ reference\ objects\ in\ a\ class}{Total\ number\ of\ reference\ objects\ in\ a\ class}$$

$$\textbf{Producer's accuracy} = 1 - omission\ error$$

III. Overall accuracy

The overall accuracy represents the proportion of reference data that has been correctly classified over all classes.

$$\textbf{Overall accuracy} = \frac{Total\ number\ of\ correctly\ classifed\ reference\ objects}{Total\ number\ of\ reference\ objects}$$

IV. Kappa coefficient

The Kappa coefficient is similar to the overall accuracy, but accounts for the possibility of classification agreement occurring by chance. It is therefore a measure of how much a classification differs from a random classification.

$$\boldsymbol{K} = \frac{Probability\ of\ correct\ classification - Probability\ of\ chance\ agreement}{1 - \ Probability\ of\ chance\ agreement}$$

2.3.7 Land cover mapping with UAV in Malta

Freely available satellite imagery (such as Landsat or Sentinel-2) typically has a low-to-medium spatial resolution (10 to 60m pixel side) which can complicate land cover mapping at large scales, for example to observe local ecological processes (Anderson & Gaston 2013). Various methods to remediate this issue have been applied successfully in Malta (Galdies et al. 2014; Gauci et al. 2018), in particular through the use of UAS (Bellia & Lanfranco 2019; Colida et al. 2017; Colida et al. 2021; Zerafa et al. 2023), however, the focus has rather been on geological processes than on land cover classification with UAS. Only Bellia & Lanfranco (2019) used unsupervised pixel-based classification of drone imagery to classify a Maltese garrigue, both at the hectare and metre scales into three land cover categories. They found that the use of a DJI Mavic 2 to classify garrigue land cover in Malta was limited by the optical resemblance between soil and bedrock covers, while the vegetation was more accurately classified. Moreover, they recommended a higher number of classes at larger scales (i.e. metre scale) to minimize the pooling of heterogeneous pixels into a single class, in particular for the vegetation.

2.4 Compositional data

2.4.1 Definitions

Compositional data is defined as "non-negative data with the distinctive property that their values [for each observation] sum to a constant, usually 1 or 100%" (Greenacre 2019, p.1). This is known as the *constant sum constraint* and it is implemented through the *closure operation*, corresponding to the division of each part by their constant sum which must be equal for each observation, resulting in a *closed* compositional dataset (Alenazi 2020). For example, in geology, when each observation corresponds to a sample with a fixed weight (e.g. 10 grams), the measured elements contained in each sample can be expressed as proportions of the total weight of a sample, and the sum of all element weights at each observation should equate the 10g constant sample weight. Similarly, when land covers are considered inside units of fixed dimensions such as grid cells, they become proportions constrained to always sum to the cell's surface. As such, inside each cell, an increase in one land cover would necessarily imply a decrease in others, biasing any correlational analysis without prior transformation.

Therefore, an observation in a composition can be represented as a positive vector $\mathbf{x}$, with D parts (i.e. variables) summing up to a constant positive κ, respecting the conditions (Filzmoser & Hron 2009) expressed by Equation 1 below:

$$\mathrm{x} = (\mathrm{x}_1, \dots, \mathrm{x}_D)^{\mathrm{t}} \quad and \quad x_i > 0 \quad and \quad i = 1, \dots, D \quad and \quad \sum_{\mathrm{i}=1}^{D} \mathrm{x}_{\mathrm{i}} = \kappa \qquad (1)$$

This set of conditions describes what have been termed by Aitchison (1986) as the simplex sample space $\mathbf{s}^D$ forming a $(D-1)$-dimensional subset of the real space $\mathbb{R}^D$. Three main properties emerge from this definition:

- Scale invariance: a change of scale in the original data (e.g. conversion from metres to centimetres) does not affect the derived compositional data as the latter carries only relative information.
- Permutation invariance: The order of parts in a composition do not affect their proportions, implying that columns of a dataset can be swapped without effects on the results, as with classic multivariate statistics.
- Subcompositional coherence: A subcomposition comprises only a subset of parts from the original composition, and any composition can usually be viewed as a subcomposition of a larger composition, such as nutrition tables on food items that do not present all nutrients but only a subset of them. A subcomposition is closed again in a process termed *reclosure* by dividing each subcompositional part with the new subcomposition sum. The results of a subcomposition should not differ from the results of the same parts in the full composition, but when using original part values, if one part is excluded and the rest of the data is closed again, the portions of the remaining parts will be different than before the exclusion. On the other hand, ratios between parts remain unchanged regardless of the exclusion of other parts and should therefore be used for analysis (Greenacre 2019).

Two main issues are often encountered when analysing untransformed compositional data with classical univariate or multivariate statistical tools, both related to the constant sum constraint described above. The first issue was pointed out by Pearson (1897) in a seminal paper warning that *spurious correlations* arise when computing correlations of ratios sharing a common denominator, as it is the case with closed compositional data. The second issue appears because each row and column of the covariance matrix of a closed compositional set has to sum up to zero due to the constant sum constraint. Therefore, as variances are always

positive, some covariances are artificially forced into negative values, thereby causing a *negative bias* in the covariance matrix (van den Boogaart & Tolosana-Delgado 2013). Both can be intuitively related to the fact that when the proportion of a part increases, the proportions of some other parts of a compositional set must decrease due to the constant sum constraint.

2.4.2 Compositional data preparation

In general, only certain parts are considered in a sample and there is a remainder which is not measured, but nonetheless required to reach the total sum of a sample and perform the closure operation. This remainder can be calculated for each observation by subtracting the sum of the measured parts from the constant total, thereby forming a new part in the composition typically called 'other'. This process is known as *completion* and prevents losing information on the total of parts, in opposition with the reclosure operation (Boogaart & Tolosana-Delgado 2013).

Ratios of compositional parts are always positive but can have a highly skewed distribution, with two standard distributions below the mean often leading to impossible negative values (Greenacre 2019). As a result, log-transformation of ratios (*log-ratios*) is a standard procedure to convert ratio-scale, positive-only numbers into real numbers as well as to linearize their ratios, enabling the use of most statistical tools based on interval scale. However, this procedure raises the issue of zeros in a compositional set which prevents the use of log-transformations.

A large number of methods have been developed to replace zeros in compositional datasets, according to the nature of zeros and the log-transformations used (Alenazi 2020; Boogaart, Tolosana-Delgado & Bren 2006; Martin-Fernandez, Barcelo-Vidal & Pawlowsky-Glahn 2003). In the compositional framework, zeros can be either *essential zeros,* the absolute absence of a

part in an observation, or *rounded zeros,* parts that are present below a detection limit for an observation (Martin-Fernandez, Barcelo-Vidal & Pawlowsky-Glahn 2003). If a part containing essential zeros has little significance for the study or if it can be related to another part, it can be amalgamated (i.e. summed) with another part before analysis. However, rounded zeros do not indicate an absence or irrelevance of the measurement but more likely a very small proportion of a part. Thus, they can be replaced with suitable low values, but the imputation method has to respect the covariance matrix to prevent changes of the structure between parts.

2.4.3 Aitchison geometry

The goal of transforming compositional data with log-ratios is to convert the simplex sample space into the real vector space where standard univariate or multivariate statistics can be applied (Greenacre 2021). Log-ratios allow to define a Euclidean vector space of the simplex by defining the *Aitchison inner product* and the *Aitchison distance* between two compositions, as well as the *Aitchison norm* of a vector (Filzmoser et al. 2018).

V. The inner product of two compositions $\mathbf{x}$ and $\mathbf{y}$ taken in $\mathbf{S}^D$ is defined as:

$$\langle \mathrm{x}, \mathrm{y} \rangle_{\mathrm{A}} = \frac{1}{2\mathrm{D}} \sum_{i=1}^{D} \sum_{j=1}^{D} \ln \frac{\mathrm{x_i}}{\mathrm{x_j}} \ln \frac{\mathrm{y_i}}{\mathrm{y_j}} \tag{2}$$

VI. The norm of a composition $\mathbf{x}$ is defined as the inner product with itself:

$$\|\mathrm{x}\|_A = \sqrt{\langle \mathrm{x}, \mathrm{x} \rangle_{\mathrm{A}}} = \sqrt{\frac{1}{2\mathrm{D}} \sum_{i=1}^{D} \sum_{j=1}^{D} \left(\ln \frac{\mathrm{x_i}}{\mathrm{x_j}} \right)^2} \tag{3}$$

VII. The distance between compositions $\mathbf{x}$ and $\mathbf{y}$ is defined as:

$$d_A(\mathrm{x},\mathrm{y}) = \sqrt{\frac{1}{2\mathrm{D}}\sum_{i=1}^{D}\sum_{j=1}^{D}\left(\ln\frac{x_i}{x_j}-\ln\frac{\mathrm{y_i}}{\mathrm{y_j}}\right)^2} \tag{4}$$

2.4.4 Log-ratio transformation

Following the definitions above, it is possible to express a composition in the Euclidean real space as coordinates in the simplex space using log-ratios transformations. There are four main log-ratios transformations (Hron et al. 2021), each with their own limitations and benefits, with the natural logarithm being traditionally used in all of them. The simplest transformation is called "pairwise log-ratios", consisting in the logarithm of the ratio of two parts from a compositional dataset. When all compositional parts are included in a subset of (J-1) linearly independent LRs, the rest of the LRs can be deduced through linear combinations. Greenacre (2021, p.274) mentions that "for a J-part composition $\{\mathrm{x_1}, \ldots, \mathrm{x_j}\}$, there are ½J(J-1) unique LRs in the form of":

$$\mathrm{LR(j,j')} = \log\left(\frac{\mathrm{x_j}}{\mathrm{x_{j'}}}\right) = \log(\mathrm{x_j}) - \log(\mathrm{x_{j'}}), \qquad j/j' = 1,\ldots,J \qquad j < j' \tag{5}$$

2.4.5 Total variability

Aitchison (1986) proposed the concept of 'total variability' in the analysis of compositional data to assess the overall variance within a compositional set. This measure focuses on the log-ratio variances computed from all unique pairs of parts in a given sample through either unweighted or weighted formulas. Weighting is required when there are parts

occurring in small amounts which can produce large components of log-ratio variance dominating the total variance. Thus, weighting serves to balance the influence of each component to ensure a more representative and unbiased calculated total variance.

For a compositional dataset with I observations ($i = 1, \ldots, \mathrm{I}$) on J parts ($j = 1, \ldots, J$) and a pair of parts denoted *jj'*, the mean of its log-ratio can be expressed as:

$$\bar{X}_{jj'} = \frac{1}{I} \sum_{i=1}^{I} \log \left(\frac{x_{ij}}{x_{ij'}}\right) \tag{10}$$

The variance of this *jj'* log-ratio is defined as:

$$\mathrm{Var}_{jj'} = \frac{1}{I} \sum_{i=1}^{I} \left(\log\left(\frac{x_{ij}}{x_{ij'}}\right) - \bar{X}_{jj'} \right)^2 \tag{11}$$

The total unweighted log-ratio variance is defined by Greenacre (2019) as all pairwise differences between log-ratios for each jj' pair of parts and each ii' pair of observations:

$$\mathrm{TotVar\ (Aitchison)} = \frac{1}{\mathrm{I}^2} \sum_{i<i'} \sum_{j<j'} \left(\log \frac{x_{ij}}{x_{ij'}} \frac{x_{i'j'}}{x_{i'j}} \right)^2 \tag{12}$$

Various weighting systems can be constructed based on knowledge of measurement errors or systematic approaches (Greenacre 2019). For varying row and columns weights, the weighted total variance is defined by Greenacre (2019) as:

$$\text{TotVar} = \sum_{i<i'} \sum_{j<j'} r_i r_{i'} c_j c_{j'} \left(\log \frac{x_{ij}}{x_{ij'}} \frac{x_{i'j'}}{x_{i'j}} \right)^2 \quad (13)$$

Equation 13 above becomes the unweighted formula when row weights $r_i = 1/\mathrm{I}$ and column weights $c_j = 1/\mathrm{J}$, after dividing it by J^2. Without expert knowledge, Greenacre (2019) reports that compositional datasets are usually weighted proportionally to their marginal totals, such as row weights $r_i = 1/\mathrm{I}$ and columns weights $c_j = j$-th part mean with $c_1 + c_2 + \ldots + c_j = 1$.

2.4.6 Variables selection

There are multiple sets of linearly independent LRs for a single compositional set of *J* parts, defined by the Cayley number J^{J-2} (Greenacre 2021). Hence, selecting a subset of LRs can be desirable if they still explain a sufficient portion of the total variability mentioned above (Hron et al. 2021). Greenacre (2019) defines the *contained variance* of a log-ratio as its proportional contribution to the total variability of a compositional set, such as $\text{ContainedVariance}_{jj'} = \text{Var}_{jj'}/J^2$. On the other hand, the *explained variance* of a log-ratio refers to how much total variability can be explained by the log-ratio while accounting for its relationships with other log-ratios, such as only "J − 1 independent LRs are needed to explain 100% of the total variance of all ½J(J − 1) LRs" (Greenacre 2021, p.278).

Greenacre (2019) mentions that Redundancy Analysis (RDA) can be used in the context of compositional data analysis to measure how much the total variability of all LRs can be explained by a subset of LRs. Given a compositional dataset encapsulated in a response matrix Z with dimensions n×m, where n is the number of samples and m is the number of parts or log-ratios, RDA employs an explanatory variable matrix W of dimensions n×p. The aim is to project Z onto the space defined by W to yield a new matrix Z*. This projection allows for the partitioning of the total variance in Z into two orthogonal components: one that is 'explained' by W and another that remains 'unexplained', corresponding to the residuals in Z-Z*.

Procrustes analysis is another method to measure the difference between the configurations of two multidimensional sets by rotating, scaling, and translating them (Legendre & Legendre 2012). It yields a *Procrustes correlation* value, bound between 0 and 1, used to measure how similar a configuration of a subset of LRs is to the configuration of the full set of LRs by comparing LRs distances in the two sets, with values closer to 1 indicating a stronger similarity (Greenacre 2018). Both RDA and Procrustes analysis are implemented together in the EasyCODA package for R (Greenacre 2018) through the *STEP* function, which provides as outputs the sequence of variables with the highest contribution to the cumulative explained variance, along with their respective R^2 and Procrustes correlation values.

2.5 Regression models

2.5.1 Linear regression

Legendre and Legendre (2012, p.536) define the purpose of regression analysis as describing "the relationship between a dependent random variable y (DV, or response) and a set of independent variables x_m (IV, also named explanatory variables or predictors), in order to describe or forecast the values of y for given values of the independent variables x_1, …, x_m". A linear regression analysis with one predictor is termed simple linear regression, while it is called multiple linear regression when there are multiple predictors. The general formulation for a multiple linear regression with i observations and m explanatory variables is:

$$y_i = \hat{y}_i + \varepsilon_i = \beta_0 + \beta_m x_{mi} + \varepsilon_i \quad (14)$$

Where $\varepsilon_i = y_i - \hat{y}_i$ is called the residual (or error) and represents the difference between the observed and estimated values of y, parameter β_0 is the value at which the regression line

intercepts the ordinate axis, and parameter β_m is the slope estimate of the regression line associated with the explanatory variable x_m, named *regression coefficient*.

A model is a general mathematical formulation describing a set of relationships between variables, with some parameters left to be estimated (Legendre & Legendre 2012). A Type I model assumes that the explanatory variables are controlled by the researcher and/or measured without error, while a Type II model assumes that these variables are random and measured with error (Legendre & Legendre 2012). In the first case, the Ordinary Least Squares (OLS) criterion is used to estimate the parameters of the model, while various other methods are used instead for Type II models, the most common being Maximum Likelihood (ML) and Major Axis (MA) estimations (Legendre & Legendre 2012). As Faraway (2022) mentions, linear regression analysis implies that the parameters β_m enter the model linearly (i.e., additively), but this does not mean that the explanatory variables themselves must be linear (e.g., they can be a logarithm or other non-linear functions). In standard regression modelling, the parameter β_m associated with an explanatory variable x_m is derived by assessing the effect of a one-unit increase in that variable on the response variable y while holding all other predictors constant (Legendre & Legendre 2012).

2.5.2 Generalized Linear Models

The linear modelling described above is in fact a special form of the Generalized Linear Models (GLMs) proposed by Nelder and Wedderburn (1972). Unlike conventional linear models that directly link the dependent variable y to the predictors, GLMs incorporate a link function, denoted as *g()* below, which allows for a more flexible relationship between the mean of the dependent variable and the linear predictors (Dunn & Smyth 2018). This relationship is mathematically represented as:

$$g(y_i) = \beta_0 + \beta_1 x_{1i} + \ldots + \beta_m x_{mi} \quad (15)$$

In this context, the variance of the response variable, usually termed θ^2, and its mean, $\mu(y)$, are used to determine the appropriate probability distribution for the model. Unlike simple linear regression, which typically assumes a constant variance, GLMs can accommodate a diverse range of variance-mean relationships by selecting suitable distributions from the Exponential Family (EF).

The EF encompasses most common probability distributions (e.g. normal or Gaussian, Poisson, Gamma, etc…), which can have their probability density function expressed as what is termed the *canonical form* of the EF (Hardin & Hilbe 2018):

$$f_y(y;\theta,\phi) = \exp\left(\frac{y\theta - b(\theta)}{a(\phi)}\right) + c(y,\phi) \quad (16)$$

Where y is the unknown data or response expressed as a function of known parameters θ, which is the natural or location parameter related to the mean, and φ, which is the scale or dispersion parameter related to the variance. The mean μ of the distribution is derived from $b(\theta)$, and the variance is a product of the dispersion parameter ϕ and a variance function $v(\mu)$. This variance function is pivotal as it describes how the variance of the distribution changes with the mean. Linear regressions usually assume a constant variance (which differs from a constant variance-mean relationship), while GLMs allow for a large range of variance-mean relationships. When Equation 16 above is instead evaluated for each known observations y_i as a function of unknown parameters θ and ϕ, it provides the likelihood function:

$$L(\theta, \phi; y_1, y_2, \dots, y_n) = \prod_{i=1}^{n} \exp\left\{\frac{y_i\theta_i - b(\theta_i)}{a(\phi)} + c(y_i, \phi)\right\} \quad (17)$$

It is often expressed as the logarithm of the likelihood function for simplicity as it removes the exponentiation and put the function on an additive scale, without altering the estimation of the parameters:

$$\mathcal{L}(\theta, \phi; y_1, y_2, \dots, y_n) = \sum_{i=1}^{n} \left\{\frac{y_i\theta_i - b(\theta_i)}{a(\phi)} + c(y_i, \phi)\right\} \quad (18)$$

By finding the values of θ and ϕ that maximize the likelihood function (Equation 17), a process known as Maximum Likelihood Estimation (MLE), it is possible to get the parameter estimates that make the observed data most probable (Zeileis, Kleiber & Jackman 2008). A GLM is therefore characterized by the choice of a distribution and a link function (Dunn & Smyth 2018), allowing it to accommodate both continuous and discrete response variables when the chosen distributions belong to the EF.

The basic distribution for modelling count data is the Poisson distribution (Zeileis, Kleiber & Jackman 2008), which assumes equal variance and mean and thus, a fixed dispersion parameter ($\varphi = 1$). If the variance is larger than the mean, it indicates overdispersion of the data and a departure from the assumptions of the Poisson distribution, leading to the selection of alternative models (Zeileis, Kleiber & Jackman 2008). A quasi-Poisson model allows to have an unrestricted dispersion parameter which is then estimated from the data along with the mean and the regression coefficients. However, the quasi-Poisson model "do not correspond to models with fully specified likelihoods" which prevents the use of traditional model

comparison methods (Zeileis, Kleiber & Jackman 2008, p.5). Another common distribution for modelling count data in ecology is the negative binomial distribution, which allows to model overdispersion by fixing $\varphi = 1$ and using an additional 'shape parameter α' instead, which is often erroneously termed 'dispersion parameter' in softwares and literature (Zeileis, Kleiber & Jackman 2008). Overdispersion is often found in biological data due to aggregation processes leading to clustering of organisms or communities, making the negative binomial distribution one of the most widely used in ecological modelling (Stoklosa, Blakey & Hui 2022). Hilbe (2011) provided a complete review of regression modelling using negative binomial distributions, in particular details relating to the different parametrizations of the mean-variance relationship in negative binomial models, with the two principal ones being 'NB1' and 'NB2'. The NB2 is the traditional negative binomial model which parametrize the variance function as $v(\mu) = \mu + \alpha\mu^2$, hence termed quadratic model, while NB1 assumes a linear variance function so that $v(\mu) = \mu(1+\alpha)$. The term α is the shape parameter and when $\alpha = 0$, the negative binomial model becomes a Poisson model as $v(\mu) = \mu$, making in fact the Poisson model a special case of the negative binomial model (Hilbe 2011).

Beside the selection of an appropriate distribution family, other modelling variations allow to account for violations of Poisson distribution assumptions such as overdispersion. The two most common are the use of random or mixed-effects models, and the use of zero-inflated or zero-truncated models (Blasco-Moreno et al. 2019). To reduce the model selection bias while determining the appropriate model structure for a dataset, Campbell (2021) recommends applying either a 'seven-step procedure' using consecutive score tests to assess dispersion and zero-inflation, or selection based on information-theoretic criteria.

To compare non-nested models using ML estimation, the *Aikaike information criterion* (AIC) has been proposed by Akaike (1973) to evaluate models considering both their fitted log-

likelihood function and their number of parameters as a measure of complexity. Let L be the maximized value of the likelihood function of a model, and k the number of parameters included in the model, the AIC is defined as:

$$AIC = -2\ln L + 2k \tag{19}$$

The model with the lowest AIC should be preferred as the optimal balance between goodness-of-fit and complexity, as the likelihood value is penalized by the number of parameters in the model (Cameron & Trivedi 2013). A modified version of the AIC allows to adapt the penalization of parameters to models with a small number of observations in order to prevent AIC from selecting models with too many parameters (i.e. overfitting). Using the same variables as in Equation 19 above, AICc (for AIC 'corrected') is defined as:

$$AICc = AIC + \frac{2k^2 + 2k}{n - k - 1} \tag{20}$$

Where n corresponds to the number of observations. It can be seen from Equation 20 that as n converges toward infinity, the extra penalty term in AICc converges toward 0 and AICc will equate AIC. Another information-theoretic criterion is known as the *Bayesian information criterion* (BIC), closely similar to AIC but with a larger penalty for added parameters. BIC is defined as:

$$BIC = k \ln n - 2\ln L \tag{21}$$

Where n corresponds to the number of observations. The BIC increases the penalty for k and integrates it with a penalty for lower observations compared to the AIC criterion. Campbell (2021) offered a thorough review of the available model selection processes and criterion, as well as the risks associated with each, notably in relation to inflated type I errors. Once a model is selected, predicted counts can be calculated by reversing the link function *g()*, which, in the case of log-link function, is done by exponentiating Equation 15 so that:

$$g(y) = \log(y)$$
$$y = \exp\left(\beta_0 + \beta_1 x_{1i} + \ldots + \beta_m x_{mi}\right) \quad (21)$$

2.5.3 Regression with compositional explanatory variables

When compositional data is used as explanatory variables, each part can only increase if the other parts also decrease in relative terms due to the constant sum constraint (Coenders & Pawlowsky-Glahn 2020). A second difficulty arises when using compositional data transformed into log-ratios, because maintaining a log-ratio constant implies that both terms of the log-ratio change by a common constant factor (Greenacre 2021). Coenders and Pawlowsky-Glahn (2020) detail the implications of these two issues on the interpretation of various log-ratio transformations. They specify that when using pairwise log-ratio, the set of J-1 selected log-ratio will have a subset that increases by a common factor while all other parts decrease by another common factor such as the total sum of the parts remains constant after back-transformation.

Aitchison and Bacon-Shone (1984) were pioneers in proposing the reformulation of a log-ratio into a log-contrast (equation 14), employing a zero-sum constraint on the effect sizes of

explanatory variables to find their values that maximize correlation with the dependent variable. With the dependent variable y, a composition with J parts and ε the residual error:

$$y = \alpha_0 + \sum_{j=1}^{D} \alpha_j \log(x_j) + \varepsilon, \qquad with \sum_{j=1}^{D} \alpha_j = 0 \tag{22}$$

When using a composition with J=4 for illustration purpose, with the arbitrary selection of (J-1) pairwise log-ratios used by Coenders and Pawlowsky-Glahn (2020) and shown in Figure 2.2 below, it leads to the following equation:

$$y = \beta_0 + \beta_1 \log\left(\frac{x_1}{x_4}\right) + \beta_2 \log\left(\frac{x_2}{x_1}\right) + \beta_3 \log\left(\frac{x_3}{x_4}\right) + \varepsilon \tag{23}$$

Equation 23 can be re-expressed as a log-contrast following equation 22 as follows:

$$\begin{aligned} y &= \beta_0 + \beta_1 \log(x_1) - \beta_1 \log(x_4) + \beta_2 \log(x_2) - \beta_2 \log(x_1) + \beta_3 \log(x_3) - \beta_3 \log(x_4) + \varepsilon \\ y &= \beta_0 + (\beta_1 - \beta_2)\log(x_1) + \beta_2 \log(x_2) + \beta_3 \log(x_3) + (-\beta_1 - \beta_3)\log(x_4) + \varepsilon \end{aligned} \tag{24}$$

An acyclic connected graph is defined by Greenacre (2019) as a graphical representation of the log-ratios of a J-parts composition, showing the connections between parts without forming a cycle, therefore containing J-1 connections. This graph can be used as an aid to interpret the relationships between parts: if a part increases, then all directly connected parts must increase by the same factor while the other parts should decrease by another common factor, usually by half of the first factor (Coenders & Pawlowsky-Glahn 2020).

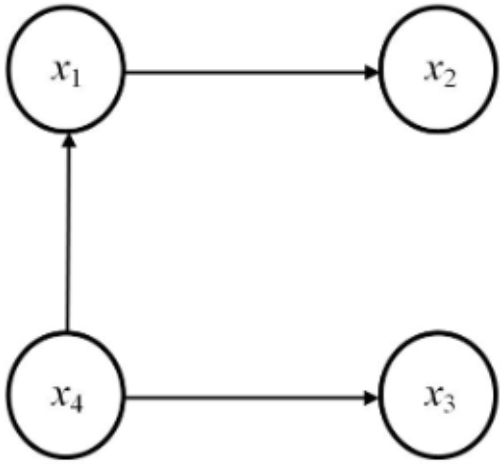


***Figure 2.3**. Example of an acyclic connected graph representing the pairwise log-ratios of a 4-parts composition, with nominators on the arrow-side of the links. From Coenders & Pawlowsky-Glahn (2020, p. 211).*

Because assessing the coefficient of each log-ratio must be done by keeping the other ratios constant, it implies that both members of the constant log-ratios change by the same factor. In this example, keeping the second and third log-ratios constant means that $\boldsymbol{x_2}$ changes by the same factor as $\boldsymbol{x_1}$, while $\boldsymbol{x_3}$ changes by the same factor as $\boldsymbol{x_4}$ in the opposite direction.

Muller et al. (2018) proposed to resort to logarithm of base 2 to facilitate the interpretation of a one-unit increase of a log-ratio, which in this case corresponds to a doubling of the ratio. Using logarithm of base 2 on the example above, the coefficient $\boldsymbol{\beta_1}$ of the first log-ratio in Equation 15 corresponds to the estimate of change in the dependent variable when doubling the proportions between the first log-ratio and the two other ones, respectively. With the previous paragraph in mind, it means that a one-unit increase in the first log-ratio will cause $\boldsymbol{x_1}$ and $\boldsymbol{x_2}$ to increase while $\boldsymbol{x_3}$ and $\boldsymbol{x_4}$ will decrease to half of the new values of $\boldsymbol{x_1}$ and $\boldsymbol{x_2}$.

2.6 Problem statement

The European rabbit (*Oryctolagus cuniculus*) is an inconspicuous yet iconic species in Malta, honouring its conservation paradox. Its role in the local culture, economy and practices is still vivid, but this is not reflected in its management or local ecological knowledge. The Maltese archipelago shares several environmental characteristics with the Mediterranean native range of the European rabbit, where it is a keystone species. While the rabbit fulfils essential ecological functions in its native range, it depends on factors that may differ significantly from the Maltese context, such as predation pressure, soil conditions or dispersal and colonisation, which raises questions on the ecological role of the wild European rabbit in Malta. In particular, elucidating its relationship with vegetation in the context of herbivory would help to determine a scientifically informed policy for both the management of natural areas and local rabbit populations. The trade-off between food and refuge availability is central to the spatial distribution of rabbits in their native range, so this research explores the relationship locally using mixed field and remote sensing techniques. Remote sensing is increasingly used in ecology and the potential of unmanned vehicles is being actively studied in Malta (Bellia & Lanfranco 2019; Colica et al. 2017). The following methodology allows to partially replace some field-based measurements and to evaluate the potential of commercial-grade drones for land cover monitoring in Malta.

3. Methodology

The availability of shelter and food are known to play an important role in *Oryctolagus cuniculus* abundance and distribution (Lombardi et al. 2003). For this dissertation, quantitative methods were used to compare the relative density of rabbit populations, assessed through a dung pellets survey, with the shrub structure, land cover composition, and anthropogenic features of 24 sites in the Maltese garrigue, assessed through a drone survey. This chapter describes the extent of the study area and the methods employed for both surveys.

3.1 Study area

The research took place on the main island of Malta inside seven areas containing an open garrigue-type vegetation community and detected signs of European rabbit activity during sites' scouting or known rabbit activity in the recent past, in addition of being physically and legally accessible as well as being outside of a drone no-fly-zone (Figure 3.1). Within each of these loosely defined Areas of Interest (AOI), three to six sites of 1 hectare each were selected according to the data collection potential left by the topography and anthropic

elements such as private property, infrastructure such as rubble walls, roads and buildings, as well as to form collectively a representative sample of each AOI (Figure 3.2).

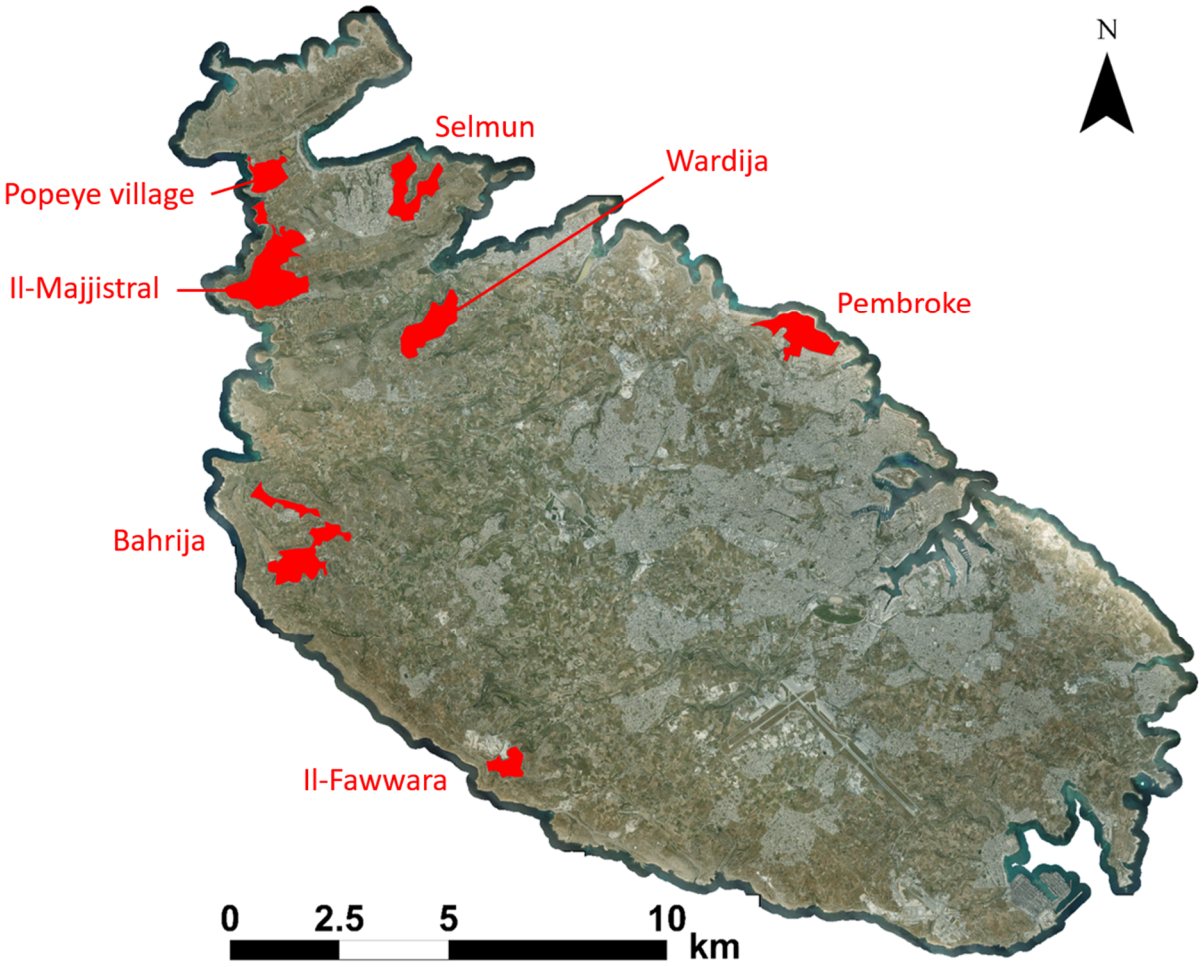


***Figure 3.1**. Location of the seven Areas of Interest (AOI) in the island of Malta (in red).*

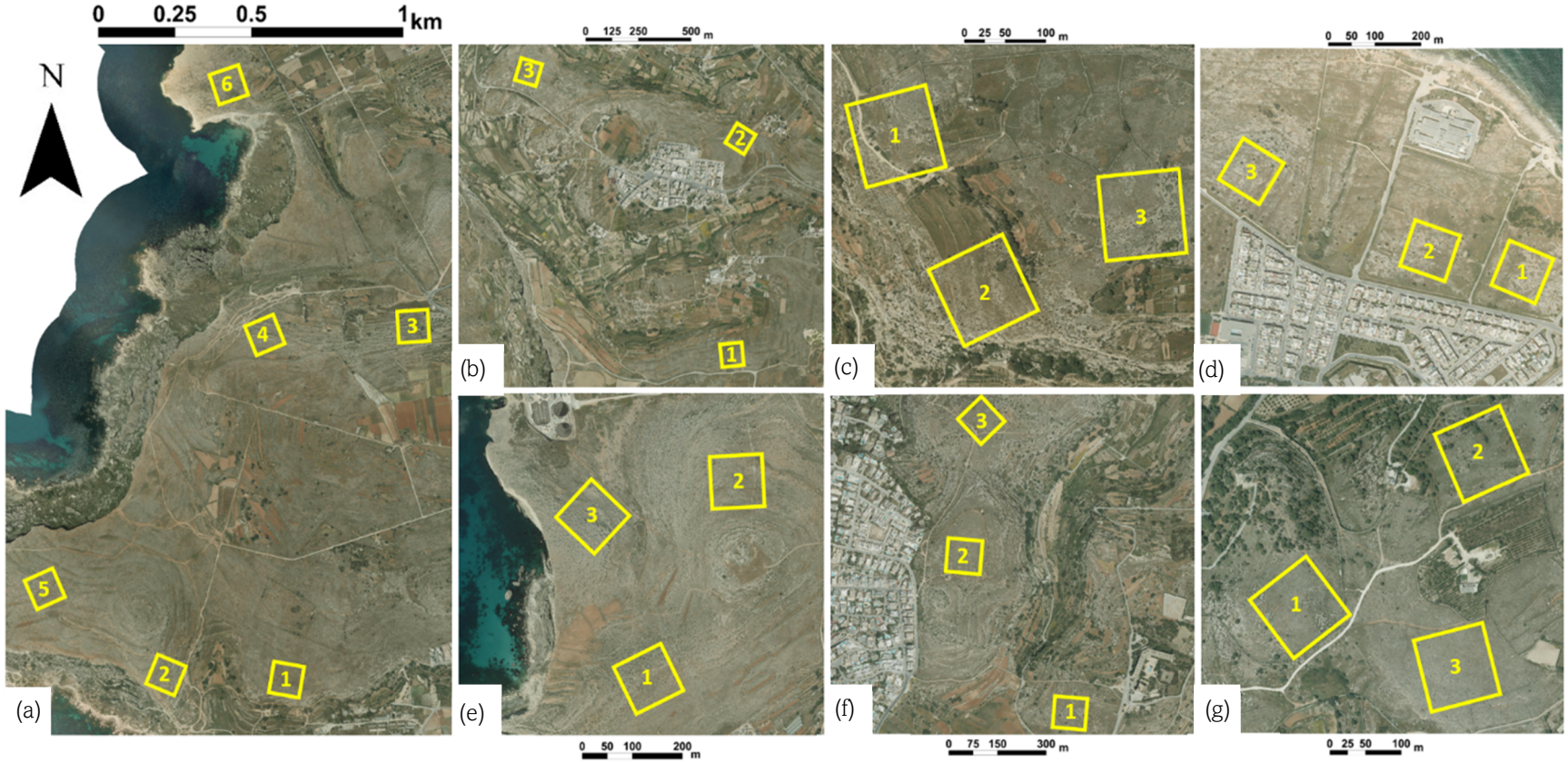


Figure 3.2. *Study sites (yellow squares, 100x100m each) inside each AOI: (a) Il-Majjistral, (b) Bahrija, (c) Il-Fawwara, (d) Pembroke, (e) Popeye village, (f) Selmun and (g) Wardija.*

3.2 Pilot study

Various combinations of hardware and software were available for the study. Therefore, a comparison of data collection workflows was carried out to optimize resources, data quality and collection repeatability (Howell et al. 2020). The first site of Il-Majjistral Park (Figure 3.2a) was used to test different sampling design parameters, the associated sampling efforts and the effects on the outputs.

Two UAS (DJI Mini 2 and Mavic 2 Pro) were flown at different altitudes and speeds calculated to obtain orthomosaics at predetermined GSD (0.5, 1 and 2 cm/pixel). The flight plans were prepared with four different methods (QGIS and Flight Planner v0.6 plug-in, Litchi v4.26.2, ArduPilot MissionPlanner v1.3.79, and Pix4Dcapture v4.8.2) to assess their workflows in terms of efforts, potential errors and output quality. Similarly, collected images were processed with two different SfM softwares (WebODM v1.9.16 and Agisoft Metashape v1.8) to compare their computing efficiency and outputs' resolution.

The Pix4Dcapture application was selected because it yielded the best match between expected and effective GSD, offered the fastest and most complete pre-flight workflow with automatic flight parameters calculations, as well as the easiest options to reliably adjust flight parameters in the field and transfer the updates back into the GIS. Unplanned adjustments in the field deplete battery levels and ultimately reduce the number of sites that can be surveyed in a single fieldwork. Agisoft Metashape software was selected for its simpler interface and lower processing time, the in-depth controls over each step of the photogrammetric process as well as for its various native algorithms, such as automated point clouds classification.

The optimal sampling intensity for European rabbit dung pellets was assessed through sequential sampling (Elzinga et al. 2001). Dung pellets were counted at the pilot site through a

systematic sampling design with a 0.1m² quadrat (n=100) placed every 10m along transects (Mutze et al. 2014, Tranchant 2022). Visible, non-degraded dung pellets were counted inside the quadrat without moving ground litter, and the mean number of pellets per quadrat, along with the standard deviation, were calculated for each new quadrat laid out. At the 100th quadrat, the running mean was 0.37 pellets/quadrat and the running standard deviation was around 1 (Figure 3.3). These values were approximately achieved around the 65th quadrat, therefore, a sampling intensity of 70 quadrats per site was selected.

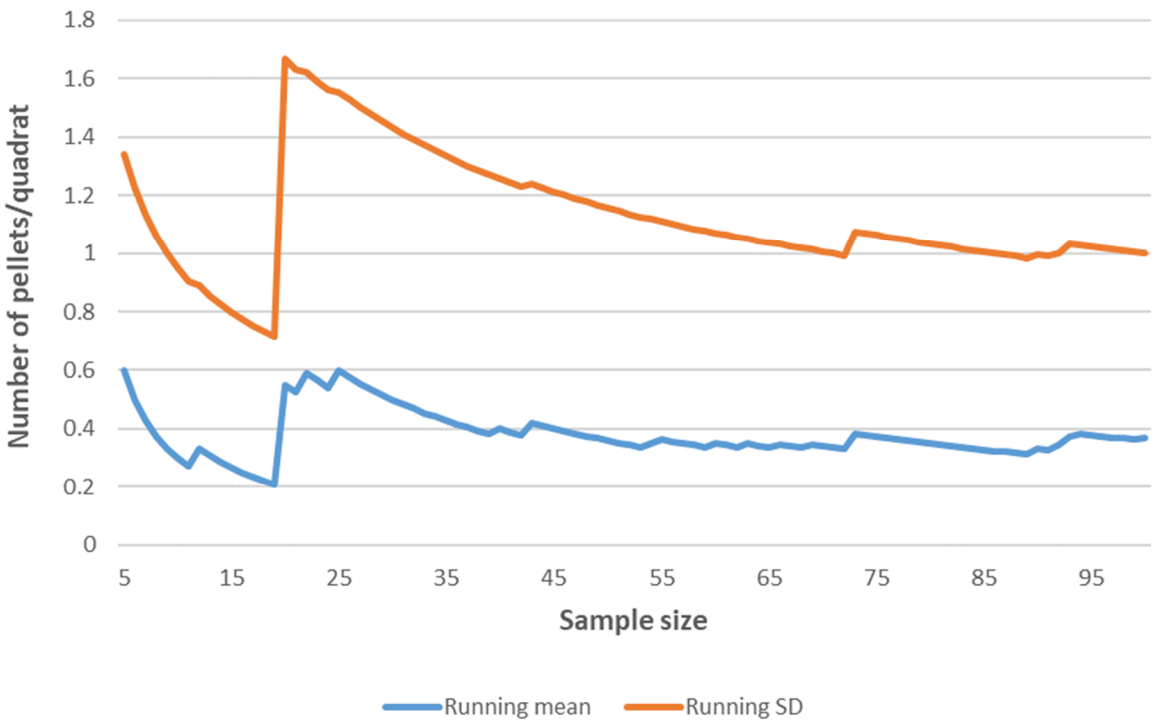


Figure 3.3. Sequential running mean (blue) and standard deviation (orange) of the number of pellets counted per quadrat according to the total number of quadrats counted.

3.3 Material & data collection

Figure 3.4 below illustrates the workflow used in this study. The data collection was undertaken in two phases: the drone survey was carried out in February 2023, while the dung pellets survey was carried out in June 2023. The drone imagery was used to create RGB orthomosaics and Above-Ground Models (AGMs), the latter allowing to accurately measure the elevation of vegetation. The AGMs were used as auxiliary inputs to the orthomosaics to enhance the classification of land covers at each site, as well as to derive various shrub measurements. The accuracy of various photogrammetric and classification procedures was compared. The land cover compositions, shrub metrics and anthropogenic features were ultimately correlated to dung pellet densities through separate generalized linear models.

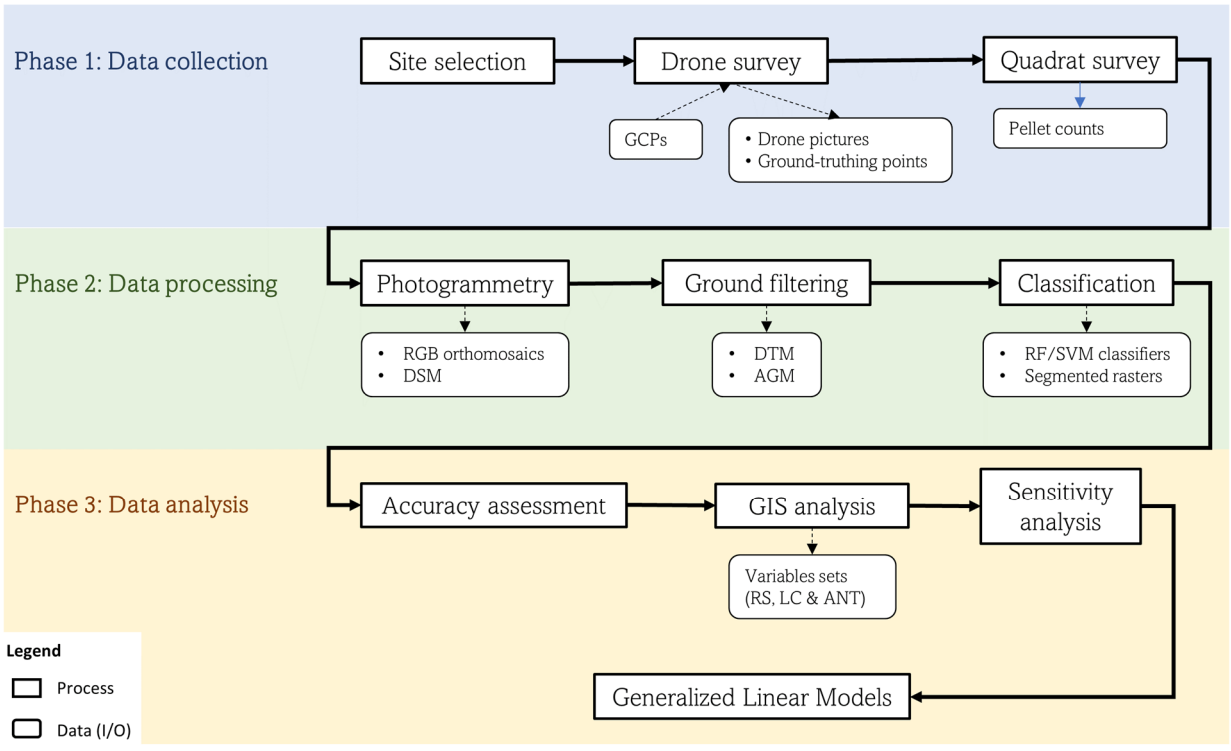


Figure 3.4. *Methodological workflow adopted in this dissertation.*

3.3.1 Drone survey

3.3.1.1 UAS

Ultra-high resolution RGB imagery (Figure 3.4) was captured during March 2023 by a DJI Mavick 2 Pro (Table 3.1) over each site between 11am and 3pm, with five GCPs placed in a cross pattern before each survey. Flights were performed as 'double grid missions' at an altitude of 40m above ground, with 80% overlap in both directions, 70 degrees' downward camera angle, and a speed of 3m/s.

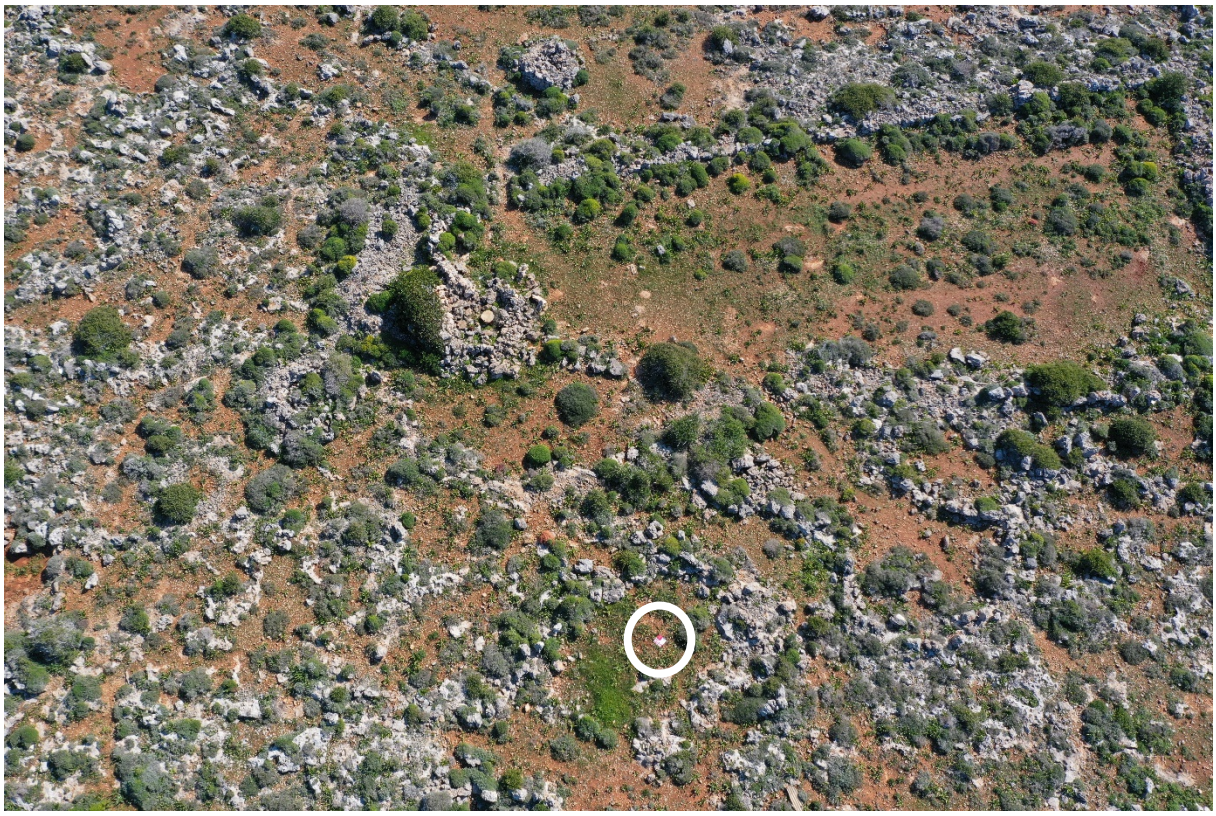


Figure 3.5 *Raw .jpg picture (RGB) from the drone survey of the land covers (40m height) at Il-Majjistral park, site 5, with one GCP visible (white circle). Typical Maltese garrigue with shrubs on soil outcrops along ridges, ruins of rubble walls and building.*

Additional GCPs were randomly acquired post-surveys in three selected sites on open ground and under shrubs for the accuracy assessment of DTMs and AGMs. The imagery captured by UAS was processed in Agisoft Metashape 1.8 software using the standard procedure for orthomosaic generation with visible-range imagery and GCPs (see Appendix 4 for detailed methodology).

Table 3.1. UAS specifications used in this study.

Characteristic	Specification
Aircraft	DJI Mavic 2 Pro
Weight (g)	907
Sensor	Hasselblad 1" CMOS (20MP)
Focal length (mm)	10.26
Image size (pixel)	5472×3648
Shutter speed	8–1/8000s
Image overlap	80% *(front & side)*
FOV (degree)	77

3.3.1.2 Comparison of ground filtering methods

Five methods of DTM generation (Table 3.2) were assessed for the accuracy of their absolute ground elevation (i.e. DTM accuracy) and vegetation height accuracy (i.e. AGM accuracy). Agisoft + CSF exhibited the highest accuracy (lower mean error), while the LiDAR method presented the greatest precision (lower standard deviation). However, the LiDAR method demonstrated superior consistency across diverse topography and terrain features, specifically concerning shrub heights, despite its lower initial spatial resolution (see Appendix 4 for details). The LiDAR approach was therefore selected.

Table 3.2. Description of the five DTM generation methods evaluated for their absolute and relative elevation accuracy.

Method name	Type of filtering	Software	Operations
CSF	Geometric	CloudCompare	Direct application of CSF tool
LiDAR	LiDAR	CloudCompare	Selection of "ground" points
Agisoft	Proprietary	Agisoft + CloudCompare	Native ground classification, cleaning and interpolation in ClouCompare
Agisoft_CSF	Proprietary + Geometric	Agisoft + CloudCompare	Native ground classification, application of CSF
ExG_CSF	Spectral + Geometric	ArcPro + CloudCompare	Filtering of ground points with Excess Green index, application of CSF

Original LiDAR point clouds from the PA were clipped for each site and imported into CloudCompare 2.13, then filtered to retain points from the 'ground' class only, from which DTMs were generated using the native Cloth Simulation Filter tool. AGMs were computed for each site, eventually corrected using a robust linear regression model developed for the occasion and applied on above-ground features only (elevation > 0) to correct the relative elevation offsets introduced by the LiDAR DTMs (Figure 3.5).

Figure 3.6. *Robust linear regression model between shrub heights derived from Above-Ground Models (x) and field measurements (y)*

Figure 3.7 below illustrates the AGM creation process, where the LiDAR DTM was subtracted from the photogrammetric DSM using a GIS raster calculator, and ultimately corrected with the above-mentioned linear regression.

Figure 3.7. The DTM derived from the PA LiDAR point cloud (top right) is subtracted from the DSM derived from the drone imagery (top left) to generate the Above Ground Model (bottom), which shows the relative elevation in meters of objects above ground, including vegetation.

3.3.1.3 Selection of a classification approach

The land covers of the Maltese garrigue landscape may have similar spectral signatures in the visible range, in particular between bedrock, soil and certain woody plant species. The type and presence of some of the target land covers were known prior to the present study (i.e. shrubs), therefore, a supervised classification approach was preferred (Richards 2022). Five classes were initially defined qualitatively to optimize spectral and elevation differences, and training samples were manually traced for each class at every site (Table 3.3, mean samples per site = 259).

Table 3.3. Initial land cover classes used in the supervised classification, with their corresponding examples and qualitative criteria, average number of training sample polygons per site, and total number of training samples.

Class name	Mean samples per site	Standard deviation	Total samples	Colour palette	Elevation
Rock *(bedrock, boulders, pebbles, artificial structures)*	79	43	1899	White, pale yellow, blue	Artifical structures only
Bare soil *(exposed soil, light gravel, pockets)*	54	27	1292	Ochre red to yellow, brown	None
Herbaceous *(grasses, annuals, forbs)*	41	30	989	Brighter green to yellow, beige	Low
High shrubs *(Medium to tall shrub species)*	31	11	735	Darker green to blue, dark red	Only with elevation
Low shrubs *(Low shrub species, e.g. Thymbra capitata)*	54	32	1296	Darker green to blue, grey and white	Very low to none

One site (Il-Fawwara 3) was selected to compare the performance of the native RF and SVM classifiers in ArcPro 2.9 over the raw RGB orthomosaic with the AGM as auxiliary input, using 200 random points previously classified manually through visual inspection of the orthomosaic only. A second site was added (Wardija 1) to compare empirically the performance of orthomosaic segmentation prior to classification using the AGM as auxiliary input and two new sets of 200 points randomly distributed over the two sites.

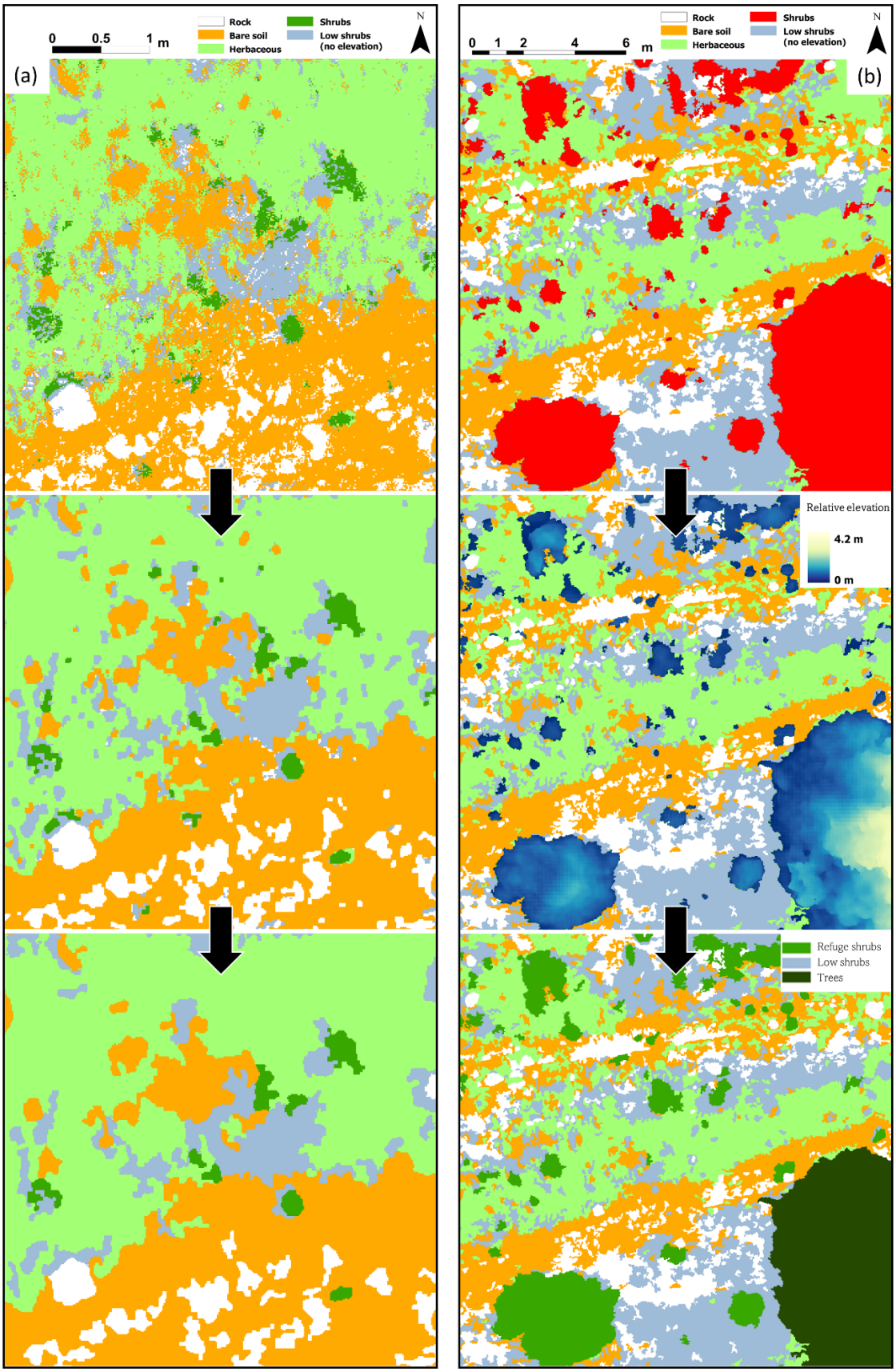


Figure 3.8. *Left column (a): illustration of the post-processing of the classified raster through generalization to remove isolated pixels. Right: illustration of the reclassification process of the initial 'shrubs' category into low-lying shrubs (<25cm), trees (>3m) and refuge shrubs (in between) land covers.*

Pixel-based SVM classifiers were trained separately for all sites using AGMs as auxiliary inputs. A post-processing generalization process (Figure 3.7) was applied to the classified rasters following the standard ESRI methodology (ESRI n.d. (a)). A second post-processing step was used to maximize the ecological relevance of the woody vegetation classes by isolating shrubs offering the highest refuge potential to the wild rabbit. The 'high shrubs' below a 0.25m threshold were considered to be more related to food availability and merged with the 'low shrubs' class, while 'high shrubs' above a 3m threshold were considered as less likely to offer a refuge potential and were thus split into a new 'trees' class. The class of remaining shrubs, more likely to be relevant for refuge availability, was renamed 'refuge shrubs'. An Average Nearest Neighbour (ANN) analysis was performed over each 'relevant shrub' to extract an average observed distance between shrub neighbours and a Nearest Neighbour (NN) index for each site (see Appendix 4 for details).

3.3.2 Dung pellets survey

European rabbit dung pellets were counted at each site with a systematic sampling design following Mutze et al. (2014) and its local adaptation by the author of this dissertation (Tranchant 2022). Eight transects per site were walked while laying down a quadrat every 12 paces (i.e. around 10m) depending on the local topography, amounting to 80 quadrats surveyed per site (Figure 3.8).

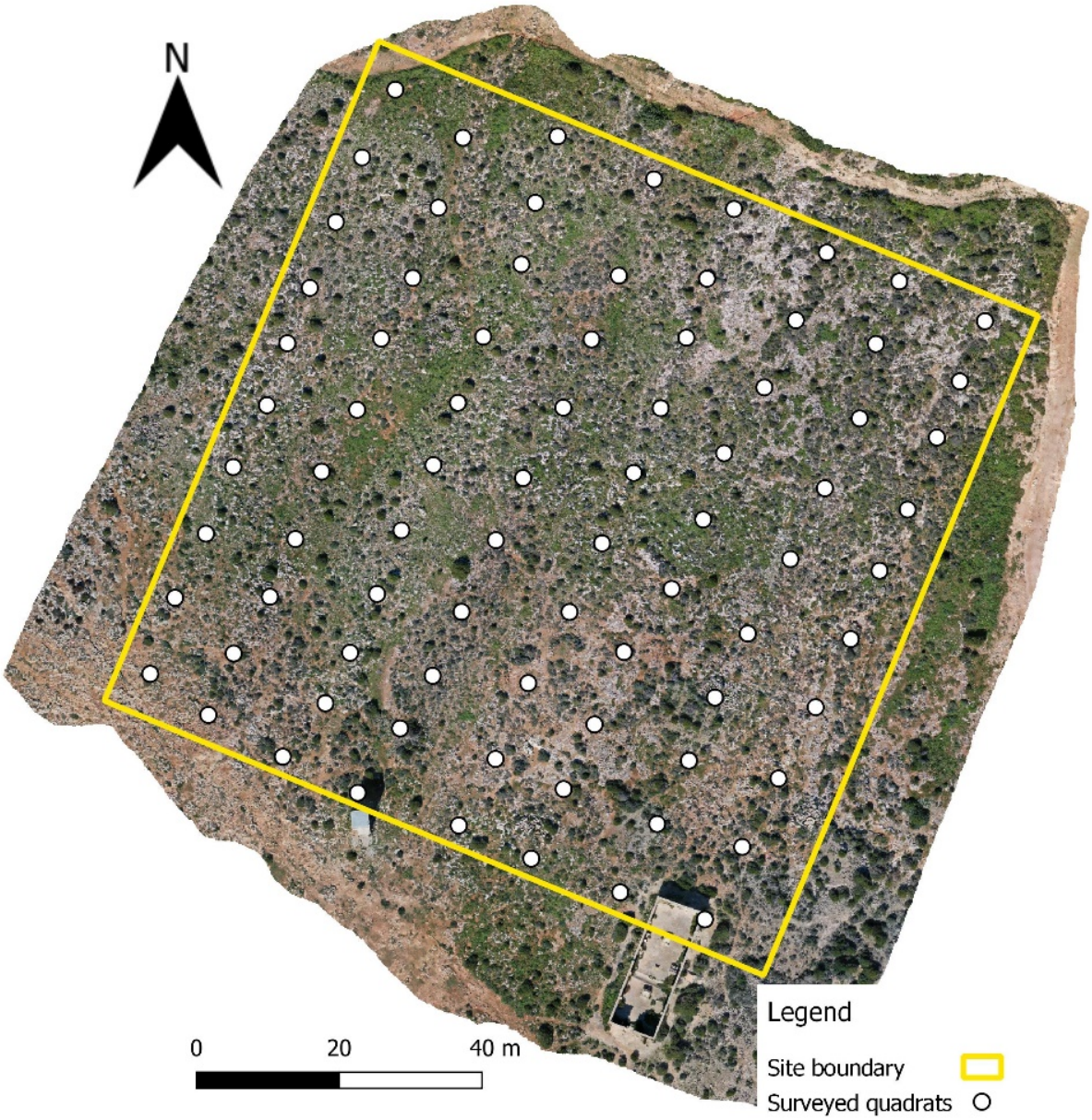


Figure 3.9. *Illustration of the systematic sampling design used in the rabbit dung pellets survey. White dots indicate quadrat locations, yellow line indicates site's extent. Rendered RGB orthomosaic of Il-Majjistral park, site 2 (1 cm/pixel).*

Non-degraded, surface dung pellets were counted inside the quadrat (Figure 3.9) and the results were logged in the LocusGIS smartphone application in real-time. The number of latrines at each site was recorded but their number of pellets not counted. Multiple factors may affect the distribution and density of dung pellets, including spatial heterogeneity induced by latrines, which should be accounted for locally when using faecal pellets to estimate absolute

European rabbit density (Palomares 2001, Putman 1984). However, because the number of faecal pellets outside of latrines correlates well with other direct and indirect density indicators, it may be used as a relative density indicator without correction for intra-seasonal comparison purpose (Iborra & Lumaret 1997). Dung pellets represent count data and was found to follow a negative binomial distribution when analysed at the quadrat level, due to the clumped deposition pattern of rabbit defecation (Tranchant 2022). Therefore, the total number of pellets counted at each site was considered as the response variable in this study to respect the count nature of pellets.

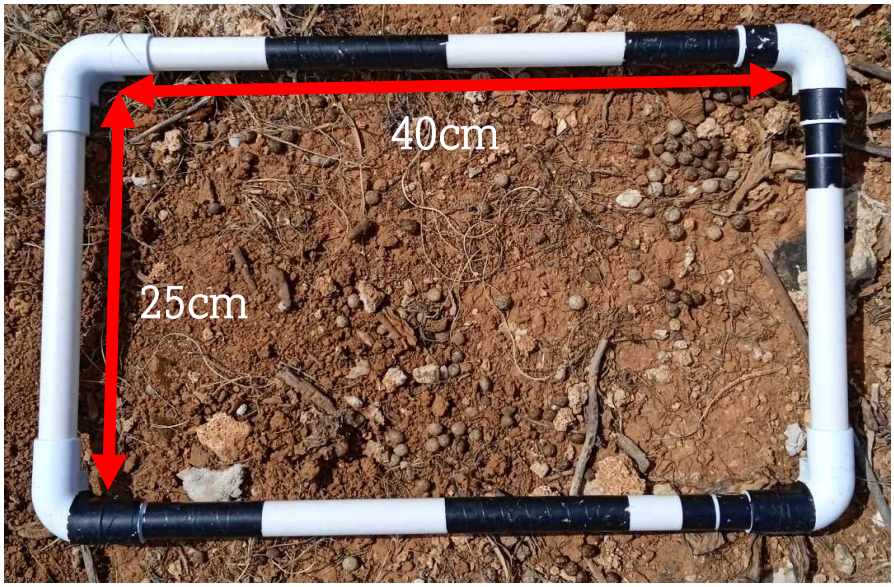


Figure 3.10. *Quadrat used for the pellet survey (0.1m²).*

3.3.3 Anthropogenic variables

Anthropogenic variables were derived from the vectorized basemap of Malta, provided by the Planning Authority of Malta, using ArcPro 2.9 and buffers of 100m, 200m and 500m around each site (Figure 3.11).

Figure 3.11. *Depiction of the 200m buffer area (blue) around Selmun's site n°2 (Selm2). Impervious roads in light grey, permeable paths in darker grey or dashed lines, light rubble walls in yellow and large rubble walls in orange. Polygon basemap from Planning Authority, Malta (2020).*

Table 3.4 below summarizes the features used in the construction of each variable. Linear features were converted to polygons of 1m width, and all features were clipped with the buffers at each scale. Subsequently, the total area occupied by the clipped features was measured in m^2 and aggregated for each variable. The roads variable corresponds to the surface of impervious roads inside buffer extents, used as an indicator of habitat fragmentation/connectivity barriers. The surface of rubble walls inside buffer extents was taken as another indicator of refuge availability, albeit that the class encompasses all rubble walls indiscriminately although their potential for refuge might vary greatly. Lastly, the surface of permeable paths, trails and tracks found inside the buffer extents was used as an indicator of human disturbance (i.e hikers, dogs, light vehicles, etc…).

Table 3.4. Parameters used to extract the three anthropogenic variables derived from the PA basemap in GIS.

Variable	Description	Layer name	Feature types
roads	Impermeable roads	GroundSurfaceCover	Road (AND material = impermeable)
rubbles	Rubble walls	OverlaidArea; OverlaidLine	Non-bounding Thin Rubble Wall, Thin Rubble Wall, Thick Rubble Wall
paths	Permeable paths	GroundSurfaceCover, OverlaidLine	Track, Country Path, Road (AND material = permeable)

3.4 Data analysis

3.4.1 Classification

Twelve sites were randomly selected to assess the accuracy of the classification results. At each site, a stratified random sampling was performed with 100 points proportionally distributed between classes according to their respective surfaces. The orthomosaics and AGMs were used to build a reference dataset for each site by manually assigning a reference class to each point through remote visual inspection, with the aid of the shrubs' GCPs. The 12 resulting confusion matrices were merged to compute an overall classification accuracy across sites. The land covers represent compositional data as the number of pixels (i.e. area in m^2) occupied by each land cover sum to the constant site surface (1ha) used in this study. The "trees" land cover class had zero values for the majority of sites, while other land cover categories did not contain zero values. The "trees" class was not accounted for in the initial study design and subsequent site selection, implying that the zero values were essential zeros, thus, the "trees" class was removed from the land cover dataset and reclosure was performed with the remaining land cover classes.

3.4.2 Statistical analysis

Data analysis was conducted using RStudio (R Core Team, 2023) with a suite of statistical packages (see Appendix 3 for full code). Pairwise LRs of variables in each compositional set were selected following the stepwise method proposed by Greenacre (2019), using the STEP function from the easyCODA package in RStudio and each component's mean as weight (Greenacre 2018). Table 3.5 below shows the sets of variables retained for the analysis. The dependent variable, 'pellets per site', was count data characterized by integers and non-negative values. A series of GLMs within the exponential family were therefore explored using a logarithmic link function on each variable set separately, with models' residual

diagnostics performed using the DHARMa package for R (Hartig 2016) and the model selection was done using AICc criterion due to the small sample size of this study. The RS and ANT sets were standardized by computing the z-scores for each variable.

Table 3.5. Definitions of the final variables retained in the analysis.

Variable set	Variable code	Variable name	Variable definition
Response	PEL	Pellets counts	Total number of dung pellets counted per site
Land covers (LC) *(m^2)*	LC1	Rock	See section 3.4.4
	LC2	Bare soil	
	LC3	Herbaceous	
	LC4	Relevant shrubs	
	LC5	Low shrubs	
Shrub metrics (RS)	RS1	Number of individual shrubs	Count of LC4 polygons
	RS2	Average individual shrub area	Mean surface of LC4 polygons (cm^2)
	RS4	Average individual maximum shrub height	Mean max. elevation of LC4 polygons (*cm*)
	RS5	Nearest Neighbour Index	Estimated / Observed inter-LC4 distance
	RS6	Observed average distance between shrubs	Average observed distance between LC4 polygons (*cm*)
Anthropogenic (ANT)	roads (100/200/500)	Impermeable roads	Surface of impermeable roads inside buffer extent (100m/200m/500m)
	rubbles (100/200/500)	Rubble walls	Surface of rubble walls inside buffer extent (100m/200m/500m)
	paths (100/200/500)	Permeable paths	Surface of permeable paths inside buffer extent (100m/200m/500m)

4. Results

In this chapter, the results of the drone survey are first presented with an emphasis on the accuracy assessment of the photogrammetry and classification procedures. This is followed by an analysis of the European rabbit dung pellet survey, where the relationship between dung pellet density and various shrub metrics is explored through a GLM approach used for inference rather than prediction. Due to the inconclusive nature of the initial findings, a second analysis is conducted, focusing on the relationship between pellet density and a range of land covers using the same modelling procedure. The chapter concludes with an examination of how pellet density is influenced by measured anthropogenic features using the same modelling framework.

4.1 Photogrammetry

The orthomosaics had a mean GSD of 1.08 cm/pixel (±0.13), which corresponds to the targeted GSD through the drone survey application, while the DSMs had a mean GSD of 2.16 cm/pixel (±0.26). The GCPs had a slightly larger vertical Root Mean Square Error (RMSE, 2.77

cm), while their total RMSE had a mean of 3.96 cm (±1.83) over all sites (Table 4.1, see Appendix 1 for details by site).

Table 4.1. Accuracy assessment of the photogrammetric processing. Values have been averaged across all surveyed sites (n=24).

		Orthomosaic						DSM	
Statistic	Flying altitude (m)	GSD (cm/pixel)	Reprojection error (pix)	GCP total RMSE (cm)	GCP X error (cm)	GCP Y error (cm)	GCP Z error (cm)	GSD (cm/pixel)	Point density (points/cm²)
Mean	50.5	1.08	0.83	3.96	1.53	1.85	2.77	2.16	0.23
Standard deviation	6.3	0.13	0.13	1.83	0.87	1.40	1.77	0.26	0.07
Min	34.9	0.77	0.60	1.30	0.52	0.44	0.76	1.54	0.14
Max	62.7	1.33	1.09	7.80	4.28	6.32	7.49	2.67	0.42

The flying altitude of the drone had a nearly perfect linear relationship (R^2=0.99) to the GSD of orthomosaics (Figure 4.1), however, there was only a range of 30m in flying height across sites, which might explain the observed linear relationship rather than the exponential one reported by Bellia and Lanfranco (2019) over a larger range of relative altitude.

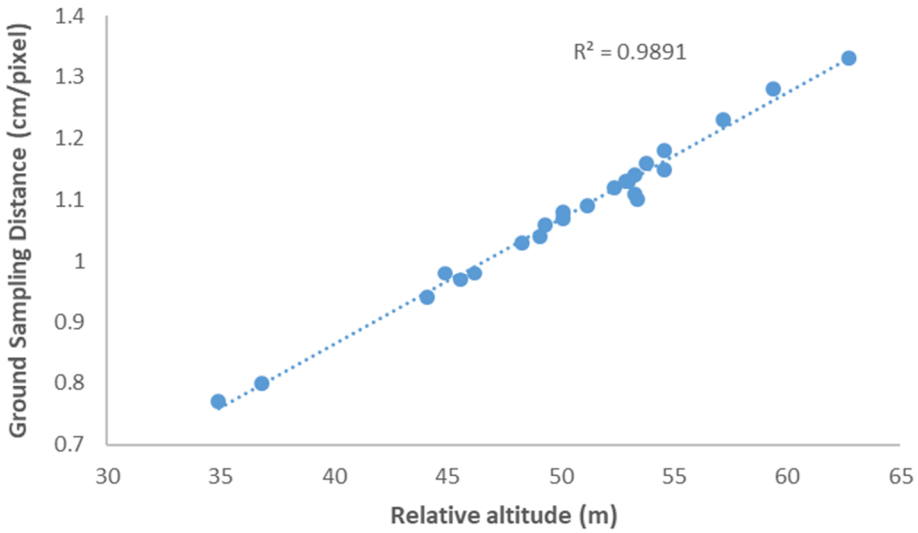


Figure 4.1 Linear regression between the flight altitude of the drone (x) and the GSD of the resulting orthomosaic (y). Blue dots are surveyed sites (n=24).

4.2 Imagery classification

4.2.1 RF and SVM classifier performance

The higher overall accuracy and kappa statistic of the SVM classifier indicated a superior performance over the RF classifier (Table 4.2), with the latter showing higher commission and omission errors for the vegetation classes. RF classifiers use a limited number of pixels (i.e. samples) for each class, here set to 1000, and increasing this number drastically reduces the computation efficiency of RF algorithm, which is one of their main advantages over SVM (Gislason 2006). On the other hand, the SVM algorithm used all pixels as samples to determine the optimal hyperplane between two classes (see section 2.3.6).

Table 4.2. Comparison of the classification accuracy using (a) the RF algorithm and (b) the SVM algorithm.

(a)		***Truth***							
	Class	Rock	Bare soil	Herbaceous	High shrubs	Low shrubs	**Total**	**User's accuracy**	**Commission error**
Predicted	Rock	62	2	2	0	2	68	91%	9%
	Bare soil	4	16	2	0	0	22	73%	27%
	Herbaceous	2	0	22	4	10	38	58%	42%
	High shrubs	2	0	6	16	10	34	47%	53%
	Low shrubs	4	2	8	0	34	48	71%	29%
	Total	74	20	40	20	56	210		
	Producer's accuracy	84%	80%	55%	80%	61%		**Overall accuracy**	**71%**
	Omission error	16%	20%	45%	20%	39%		**Kappa coef.**	**0.63**

(b)		***Truth***							
	Class	Rock	Bare soil	Herbaceous	High shrubs	Low shrubs	**Total**	**User's accuracy**	**Commission error**
Predicted	Rock	32	2	2	0	0	66	94%	6%
	Bare soil	4	18	2	0	0	24	75%	25%
	Herbaceous	4	0	32	4	6	46	70%	30%
	High shrubs	0	0	0	16	4	20	80%	20%
	Low shrubs	4	0	4	0	46	54	85%	15%
	Total	74	20	40	20	56	210		
	Producer's accuracy	84%	90%	80%	80%	82%		**Overall accuracy**	**82%**
	Omission error	16%	10%	20%	20%	18%		**Kappa coef.**	**0.77**

4.2.2 Segmentation performance

The single scale segmentation approach did not improve the classification accuracy compared to the pixel-based approach (Table 4.3). Various segmentation parameters were tested empirically in ArcPro 2.9 using the Segment Mean Shift function but could not be compared systematically due to processing time limitations. The best results were obtained with spectral detail 15, spatial detail 8, and minimum segment size 80 pixels (corresponding to ≈80cm²). Table 4.3 shows that at both sites, the object-based approach with the selected segmentation parameters had larger commission and omission errors, mainly in the vegetation classes.

Table 4.3 Comparison of the classification accuracy using (greyed) unsegmented orthomosaics and (white background) segmented orthomosaics at two selected sites.

	Wardija 1				Il-Fawwara 3			
	Unsegmented		Segmented		Unsegmented		Segmented	
Class	**Commission error**	**Omission error**	**Commission error**	**Omission error**	**Commission error**	**Omission error**	**Commission error**	**Omission error**
Bedrock	19%	19%	22%	13%	6%	16%	14%	14%
Bare soil	10%	25%	22%	42%	25%	10%	0%	70%
Herbaceous	0%	66%	48%	41%	30%	20%	43%	20%
High shrubs	5%	0%	41%	11%	20%	20%	60%	0%
Low shrubs	32%	4%	33%	51%	15%	18%	42%	75%
Overall accuracy	78%		63%		83%		65%	
Kappa statistic	0.70		0.52		0.77		0.54	

4.2.3 Classification accuracy assessment

The final image classification process, involving a pixel-based SVM classification with ultra-high resolution RGB imagery, relative elevation data and post-processing, was generally successful in distinguishing six land cover classes even in terrain with complex geomorphology such as terraced slopes as shown in Figure 4.2 below.

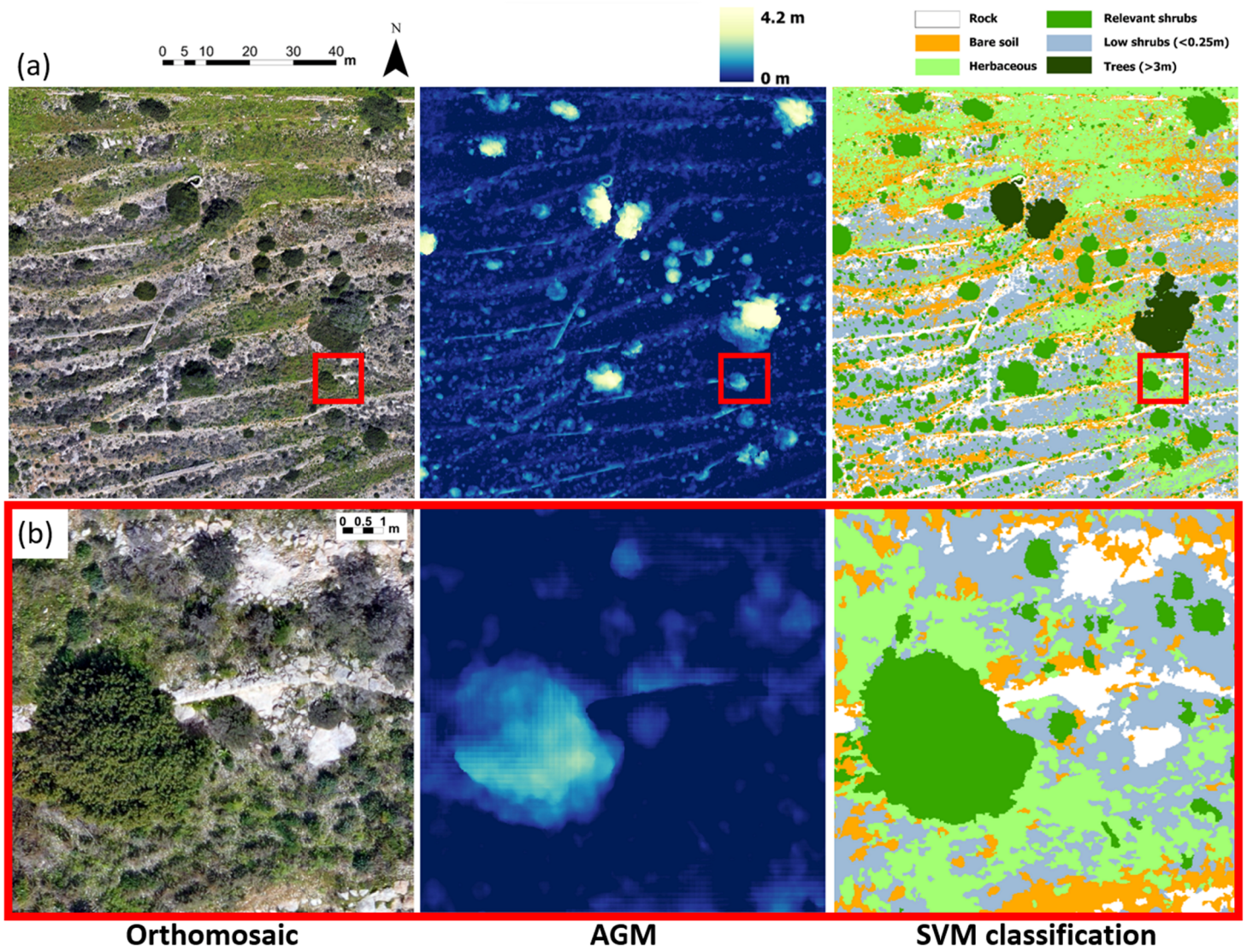


Figure 4.2. Illustration of the classification workflow using the RGB orthomosaic (left column) and the AGM (middle column) as auxiliary input, at (a) the site level and (b) the plant level. Red box in (a) indicate location of the image directly found below in row (b).

The two-steps post-processing improved the overall accuracy of the SVM classification by approximately 10%, positively affecting all classes except the 'high/relevant shrubs' class, which lost 2% of producer's accuracy but gained 7% of user's accuracy (Table 4.4).

Table 4.4. Comparison of the classification accuracy between before (a) and after (b) the two steps of post-processing. First step involved generalization while the second involved shrubs reclassification.

(a)

		Truth							
	Class	Rock	Bare soil	Herbaceous	High shrubs	Low shrubs	**Total**	**User's accuracy**	**Commission error**
Predicted	Rock	176	7	4	4	24	215	82%	18%
	Bare soil	15	135	16	0	5	171	79%	21%
	Herbaceous	4	7	227	5	27	270	84%	16%
	High shrubs	5	0	15	168	17	205	82%	18%
	Low shrubs	36	28	57	5	303	429	71%	29%
	Total	236	177	319	182	376	n=1290		
	Producer's accuracy	75%	76%	71%	92%	81%		**Overall accuracy**	**78%**
	Omission error	25%	24%	29%	8%	19%		**Kappa statistic**	**0.72**

(b)

		Truth								
	Class	Rock	Bare soil	Herbaceous	Relevant shrubs	Low shrubs	Trees	**Total**	**User's accuracy**	**Commission error**
Predicted	Rock	189	2	0	4	8	0	203	93%	7%
	Bare soil	9	157	5	0	0	0	171	92%	8%
	Herbaceous	1	3	249	3	9	0	265	94%	6%
	Relevant shrubs	6	0	8	124	1	0	139	89%	11%
	Low shrubs	31	15	55	4	358	0	463	77%	23%
	Trees	0	0	2	3	1	44	50	88%	12%
	Total	236	177	319	138	377	44	n=1291		
	Producer's accuracy	80%	89%	78%	90%	95%	100%		**Overall accuracy**	**87%**
	Omission error	20%	11%	22%	10%	5%	0%		**Kappa statistic**	**0.83**

The 'low shrubs' category had the lowest user's accuracy with 77% of its classified points corresponding to the reference data, while its producer's accuracy was high (95%), indicating that the SVM classification was successful at correctly classifying low shrubs present in the field but included other land covers as well in that category (Table 4.4b). In fact, the commission errors came from the misclassification of herbaceous (55), rock (31) and bare soil (15) as low shrubs. Accordingly, the 'herbaceous' class had the lowest producer's accuracy, with 78% of its reference points correctly classified as such, while its user's accuracy remained

high (94%). This was mostly related to the misclassification of 'herbaceous' cover into 'low shrubs' as mentioned above, with few other misclassifications into 'relevant shrubs' (8) and 'bare soil' (5).

4.3 Rabbit relative density and refuge shrubs

4.3.1 Descriptive statistics

In total, 1902 quadrats and 2200 European rabbit dung pellets were counted, excluding those found in latrines (Appendix A1). At the AOI level, Selmun had the highest mean pellet density, while the Bahrija AOI had one site with zero pellets counted and Pembroke was the only AOI without any observed dung pellets (Figure 4.3). Given that rabbit dung pellets can persist for several months (see section 2.2.3), the total absence of pellets in Pembroke suggests an absence of rabbits in this AOI at least for the observed year (2023). The current study design explicitly targeted areas with averred rabbit presence to compare changes in relative rabbit density according to vegetation. Therefore, the three Pembroke sites were excluded from the rest of the analysis.

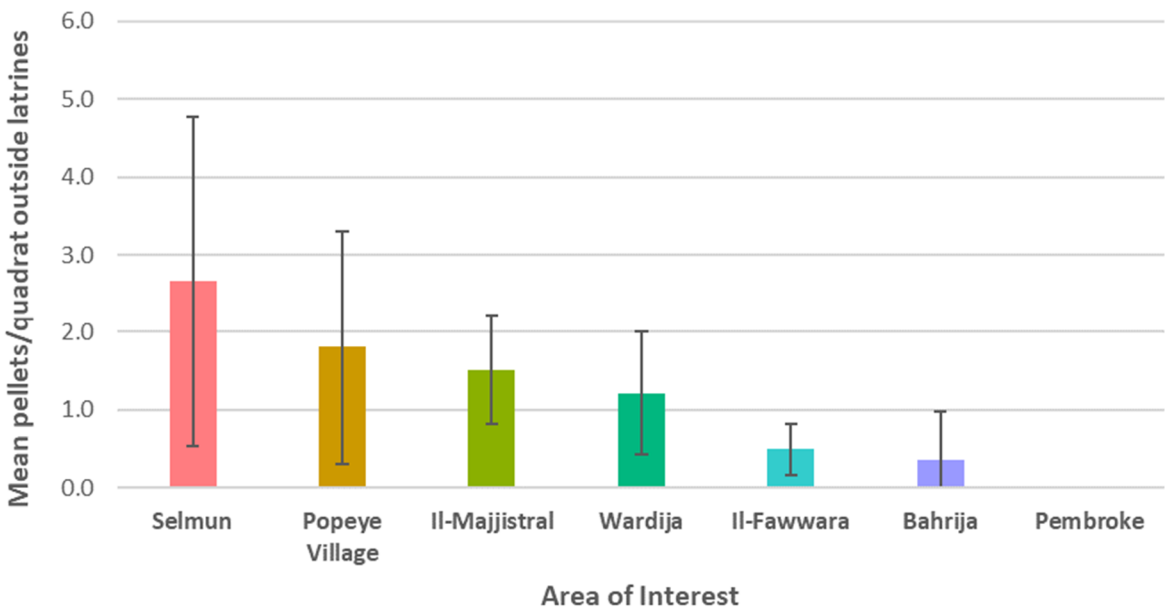


Figure 4.3. *Average pellet density (mean pellets/quadrat) at each surveyed Area of Interest (AOI).*

The total number of pellets varied between 0 and 358 counts across the 21 remaining sites, with an average of 104.8 pellets counted per site (Table 4.5). The skewness (0.97) and the kurtosis (3.4) indicated a right-skewed and slightly peaked distribution. A low proportion of null total counts was observed (4.8%), while the variance-to-mean ratio (85.9) suggested a high level of overdispersion. The overall relative rabbit density observed here was 11.57 (± 11.73) pellets/m² across all sites, which can be considered as low by international standards, with reported pellet densities of up to over 200 pellets/m² in the Iberian Peninsula (Fernandez-de-Simon et al. 2011; Rueda et al. 2008). As a comparison, a previous dissertation by the author found pellet densities around four times higher (40 pellets/m²) on Comino than those reported here for the main island of Malta (Tranchant 2022).

Table 4.5 *Descriptive statistics of the dependent variable, pellet counts per site, for all sites except the three Pembroke sites (n=21).*

Mean	104.8
Minimum	0
Maximum	358
Standard deviation	94.9
Variance	9005.4
Skewness	0.97
Kurtosis	3.4
Proportion of zeros	4.8%

Descriptive statistics for the shrub variables averaged across all sites are presented in Table 4.6 below (see Appendix 1 for details by site). The two height measurements showed a very low standard deviation, possibly due to the AGM correction method inducing a constant minimum and to the reclassification using height thresholds, which both reduced the variance. The majority of sites showed clustering in their shrub distribution pattern (NN index < 1), while three sites had a dispersed pattern (NN index > 1) and two sites had a random pattern (NN index = 1). The average observed distance to the nearest shrub neighbour showed a large variance, with values ranging from 0.28m to 1.47m.

Table 4.6 *Descriptive statistics for the shrub variables averaged over all sites beside Pembroke (n=21).*

Statistic	Number of shrubs	Average shrub area (m^2)	Average shrub maximum height (m)	Average shrub height (m)	Nearest Neighbour index	Average distance to nearest neighbour (m)
Mean	1621	0.62	0.41	0.30	0.88	0.78
Standard deviation	838	0.13	0.07	0.06	0.12	0.28
Minimum value	295	0.31	0.32	0.22	0.67	0.28
Maximum value	3304	0.90	0.57	0.45	1.07	1.47

4.3.2 Model

An initial GLM with a Poisson distribution was fitted with the pellet counts per site as the dependent variable and all the standardized RS variables as predictors. DHARMa residual tests showed significant overdispersion (dispersion parameter = 64.68, p-value < $2.2e^{-16}$) and zero-inflation (ratio = inf., p-value < $2.2e^{-16}$), while the residual plots additionally indicated a significant departure of the residuals from a normal distribution (Kolmogorov-Smirnov test p-value = 0.00451) and the presence of significant outliers (Figure 4.4). The low number of observations prevented the calculation of the quantile regressions for residual vs. predicted values.

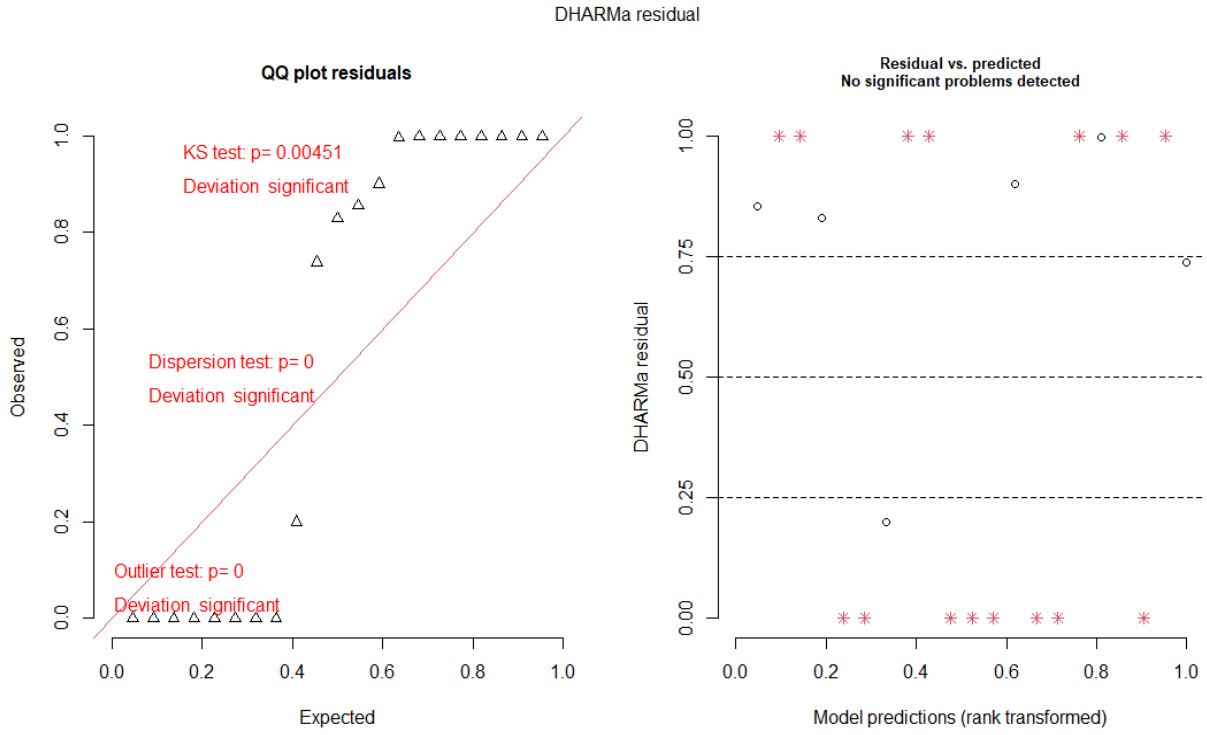


Figure 4.4 *Left: Expected (x) versus observed (y) quantile residuals for the initial Poisson GLM between pellet counts and the shrubs variables. Right: Rank transformed predicted values (x) versus residual errors (y) of the Poisson GLM.*

The large overdispersion initially indicated by the mean-variance ratio already suggested the use of a negative binomial distribution to accommodate it. To test this, the response variable was binned into five groups of equal width and the mean of each group was plotted against the respective mean squared residuals of the Poisson model (Figure 4.5). The quadratic negative binomial (NB2) was confirmed as the most adapted parametrization for this dataset (see section 2.5).

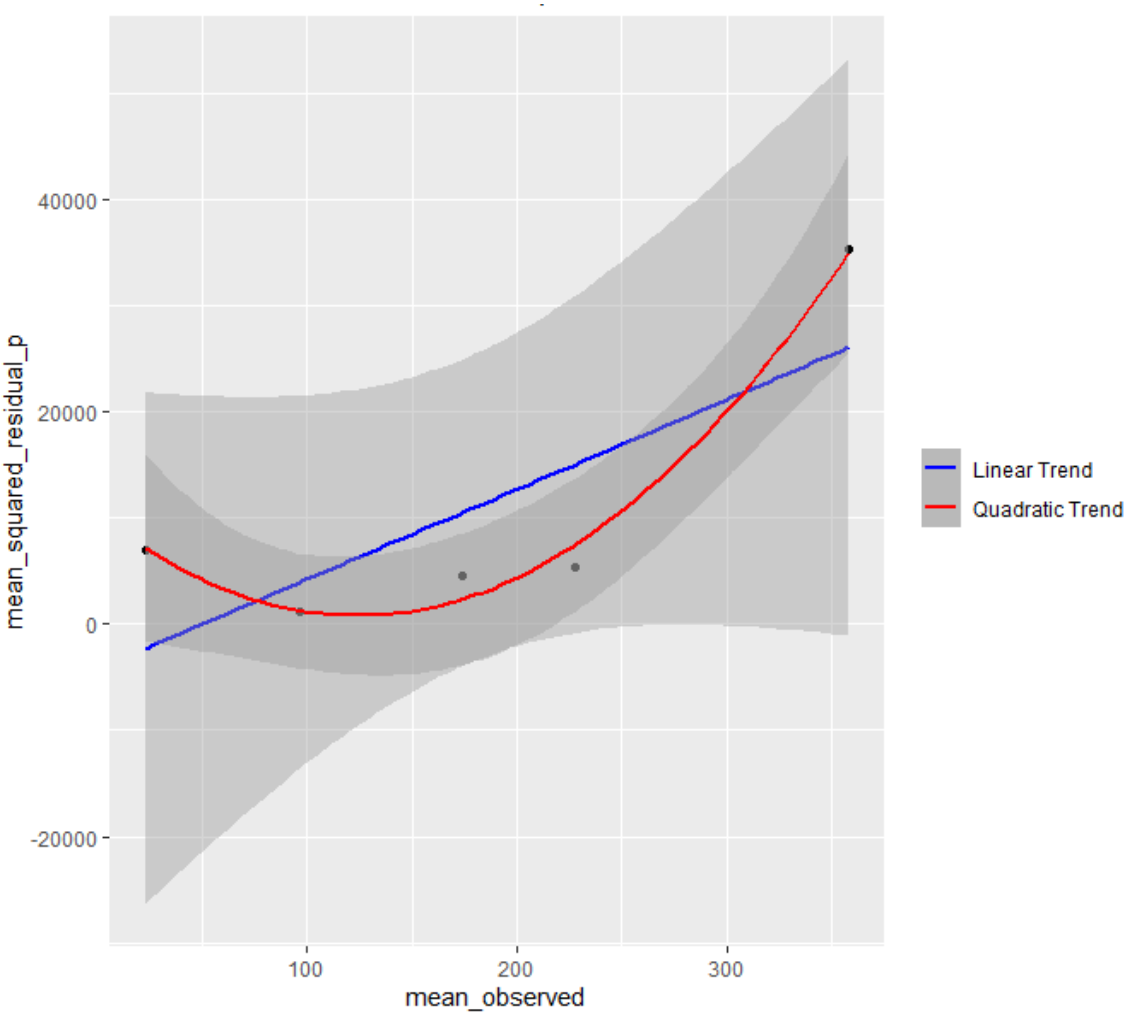


Figure 4.5 *Mean observed pellet count (x) against mean squared residuals (y) of the Poisson GLM between pellet counts (binned by equal group width) and the shrubs metrics used as explanatory variables. The blue line indicates a regression using the NB1 formula, while the red line indicates a regression using the NB2 formula.*

Consequently, a NB2 GLM (see section 2.5.2) was fitted using all RS variables as predictors. DHARMa residual tests were run with n=500 simulated observations generated from the fitted model. The Q-Q plot of the residuals showed slighter higher residual values than expected without significant departures from normality, dispersion, and outlier assumptions (Figure 4.6). The dispersion test indicated insignificant under-dispersion (dispersion value=0.55, p-value=0.56), while the zero-inflation test suggested a higher occurrence of zeros in the dataset than expected, without being significant (ratio=3.07, p-value=0.57).

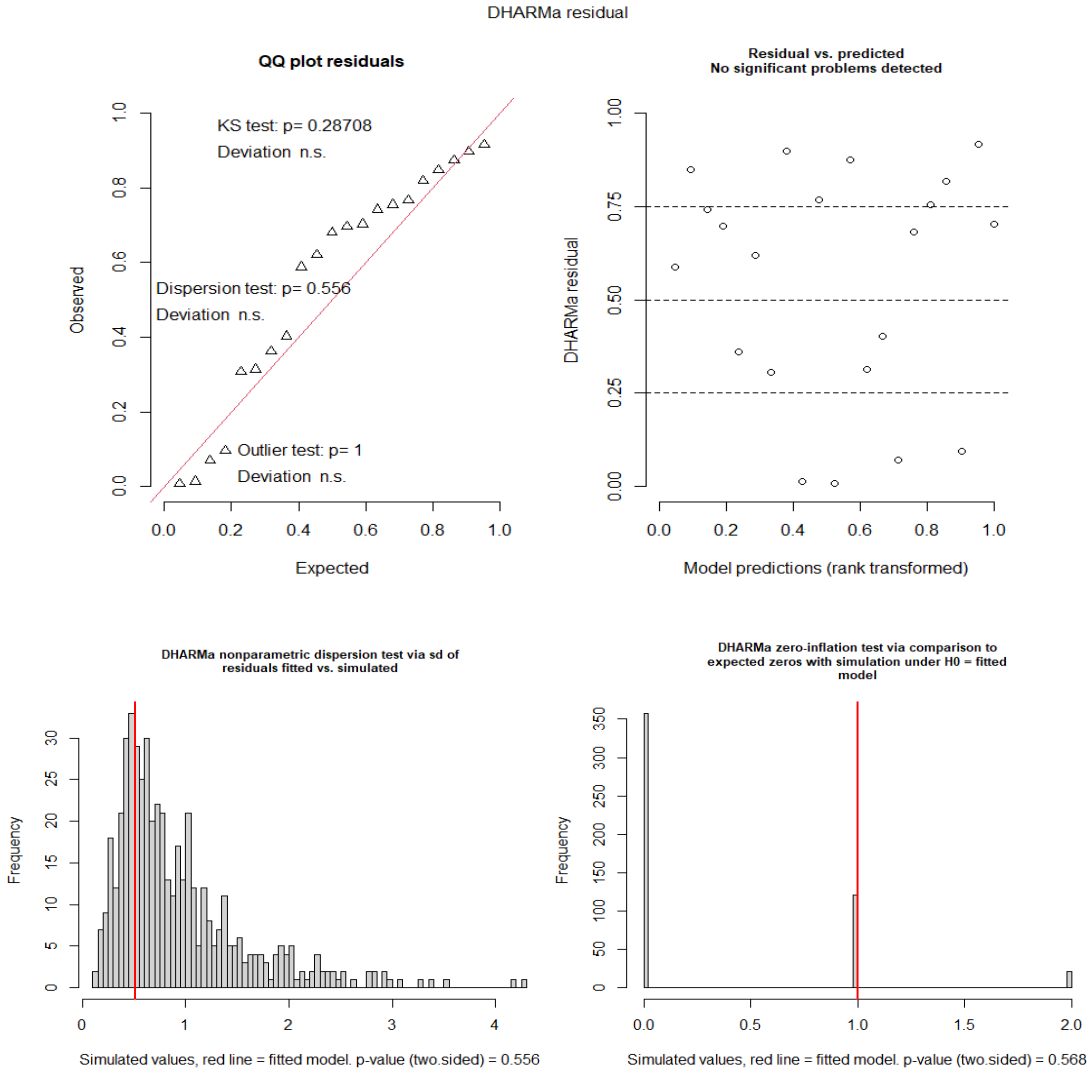


Figure 4.6. *Model validation for the NB2 GLM with pellet counts and shrubs variables using DHaRMA package for R.*

NB2 GLM with one, two, three and four predictors were compared to the null and full models using AICc criterion. Table 4.7 below presents the models with AICc weights > 0.05 and shows that the null model performed best among those compared.

Table 4.7. Selection of the NB2 GLMs with lowest AICc and AICc weight > 0.05, for pellet counts (y) and individual shrub metrics (x).

Model formula	K	AICc	Delta AICc	Model Likelihood	AICc weight	Log-Likelihood	Cumulative weight
pellets_counts ~ 1	2	241.67	0.00	1.00	0.30	-118.50	0.30
pellets_counts ~ RS4	3	243.94	2.27	0.32	0.10	-118.27	0.40
pellets_counts ~ RS5	3	244.09	2.42	0.30	0.09	-118.34	0.49
pellets_counts ~ RS6	3	244.15	2.48	0.29	0.09	-118.37	0.58
pellets_counts ~ RS2	3	244.31	2.64	0.27	0.08	-118.45	0.66
pellets_counts ~ RS1	3	244.41	2.74	0.25	0.08	-118.50	0.73

The GLM analysis for this section indicated no significant relationships between shrub metrics and rabbit pellet counts, providing insufficient evidence to reject the null hypothesis of the first research question of this thesis.

4.4 Rabbit relative density and land covers

4.4.1 Descriptive statistics

The compositional nature of the land cover dataset is illustrated by Figure 4.7 below, with sites organized by increasing dung pellets density along the vertical axis. The proportion of herbaceous cover appeared to be lower at sites with higher pellet density, while an opposite trend could be observed for the low shrubs cover with some exceptions.

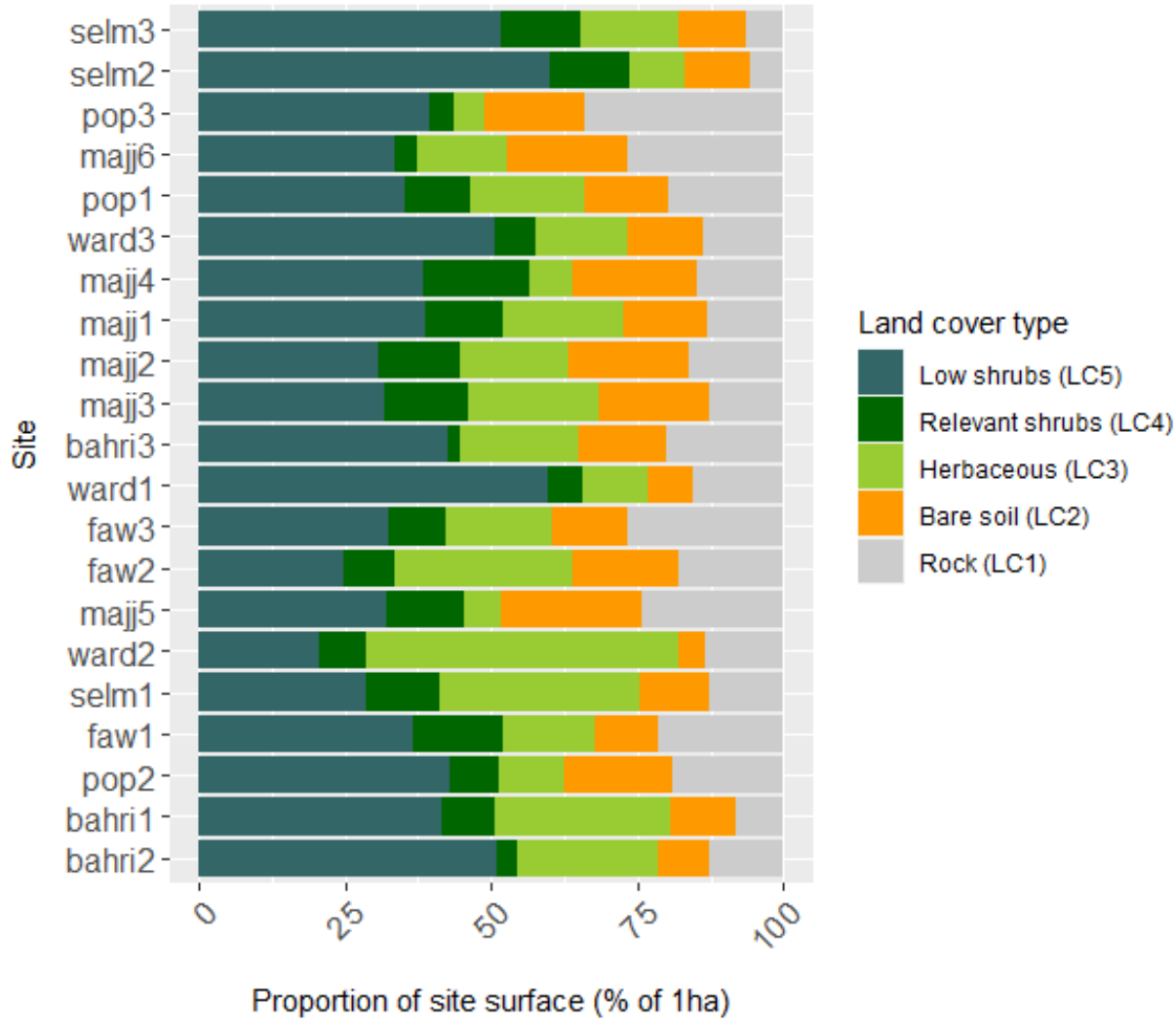


Figure 4.7. Stacked bars plot of the land cover composition at each considered sites (n=21).

However, ratios of proportions are better suited to analyse changes in compositional data due to their relative nature (see section 2.4), which can mask significant changes of ratios between parts that are difficult to detect visually such as with Figure 4.7. Hence, a stepwise selection of log-ratios was performed, and Table 4.8 below shows the set of (J-1) selected log-ratios accounting for the total explained variance of the land covers compositions.

Table 4.8. *Final selection of log-ratios (base 2) for the land cover dataset, using the STEP function in easyCODA package (Greenacre 2021).*

Log-ratio	% variance explained (R^2)	% variance explained (accum.)	Procrustes correlation (accum.)	Log-ratio interpretation
LC1/LC3	40.5%	40.5%	0.64	Ratio of rock to herbaceous covers
LC3/LC5	29.5%	70.0%	0.80	Ratio of herbaceous to low shrubs covers
LC4/LC5	22.7%	92.7%	0.92	Ratio of refuge shrubs to low shrubs covers
LC2/LC5	7.3%	100.0%	0.97	Ratio of bare soil to low shrubs covers

Additionally, an acyclic connected graph was constructed to facilitate the subsequent interpretation of the results (Figure 4.8).

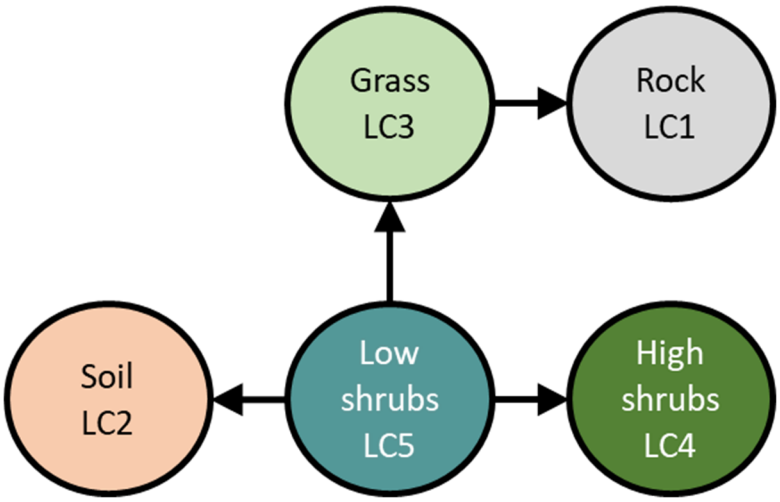


Figure 4.8. *Acyclic connected graph depicting the selected log-ratios for the land cover dataset and their relative links. The numerator of a log-ratio is found on the arrow side of a given link, while its denominator is on the non-arrow side. Adapted from Greenacre (2019).*

For the land covers dataset, the first selected LR was the ratio of rock to herbaceous covers, which described 40.5% of the total explained variance with a 0.64 Procrustes correlation value with the original dimensions of the data, suggesting that the variations in land covers of the selected sites were mainly driven by the proportion of rock cover to the herbaceous layer. The second selected LR was the proportion of herbaceous cover to low-lying shrubs, such as *Thymbra capitata*, explaining 29.5% of the total variance. The third selected LR described 22.7% of the total variance and represents the ratio of refuge shrubs to low-lying shrubs, while the last selected LR was the ratio of bare soil to low-lying shrubs, explaining the remaining 7.3% of the total variance. As shown by Figure 4.8 above, the low-lying shrubs cover was central to the variations in land cover composition of the observed sites, seconded by the herbaceous cover.

4.4.2 Model

The selected LRs were first converted to logarithm of base 2 for easier interpretation (Muller et al. 2018), then used as predictors in an initial Poisson GLM, which showed significant dispersion, non-uniformity of the residuals and zero-inflation. Consequently, a mean-variance binned plot was created for the Poisson model to select the most adapted mean-variance relationship (Figure 4.9). The mean-variance relationship was similar to the RS dataset model and the quadratic negative binomial mean-variance relationship (NB2) was selected.

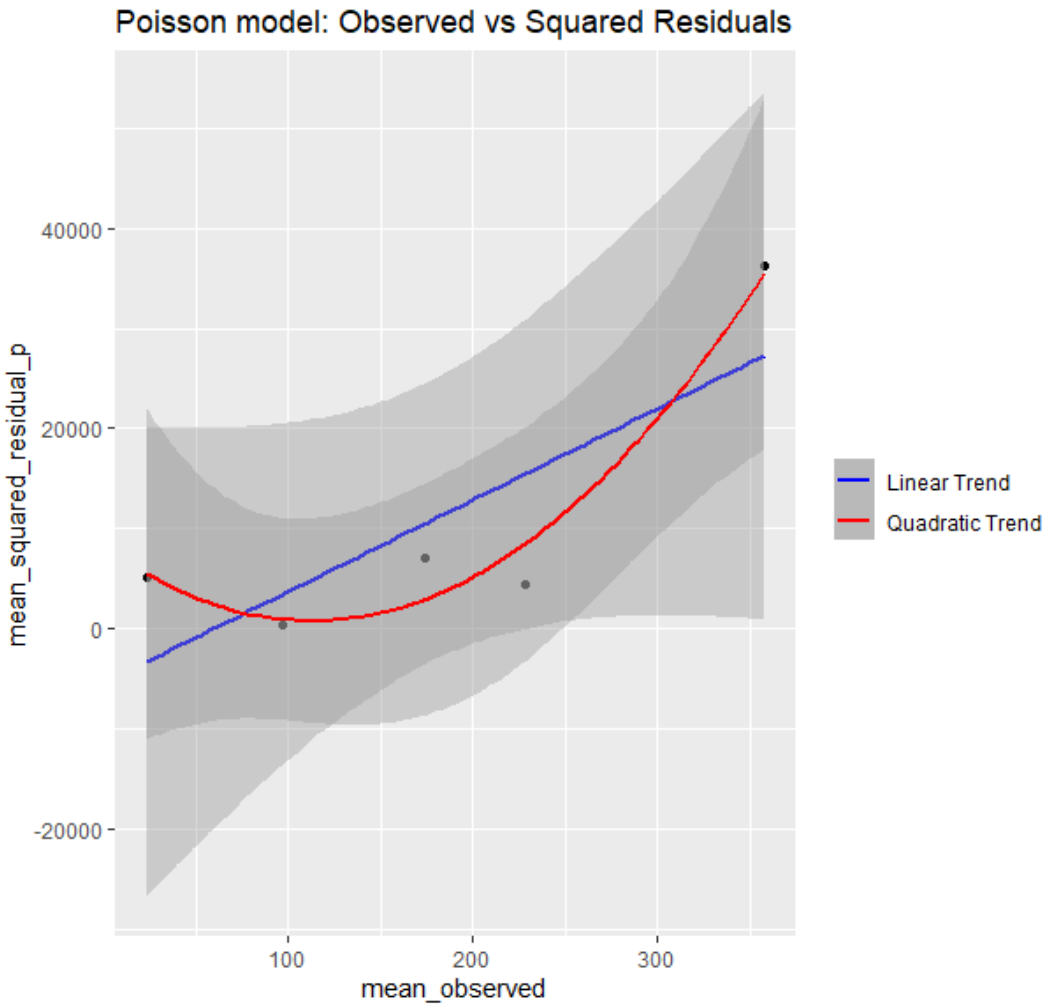


Figure 4.9. *Mean observed pellet count (x) against mean squared residuals (y) of the Poisson GLM between pellet counts (binned by equal group width) and the land cover log-ratios as explanatory variables. The blue line indicates a regression using the NB1 formula, while the red line indicates a regression using the NB2 formula.*

The residuals of the NB2 model with the four selected LRs as predictors and n=500 simulated samples appeared to be relatively normally distributed, with no significant proof of dispersion nor outliers (Figure 4.10). The ratio of expected versus observed zeros under simulation from the fitted model was larger than normal but insignificant (ratio = 3.268, p-value=0.548), while there was insignificant underdispersion (0.47, p-value=0.44).

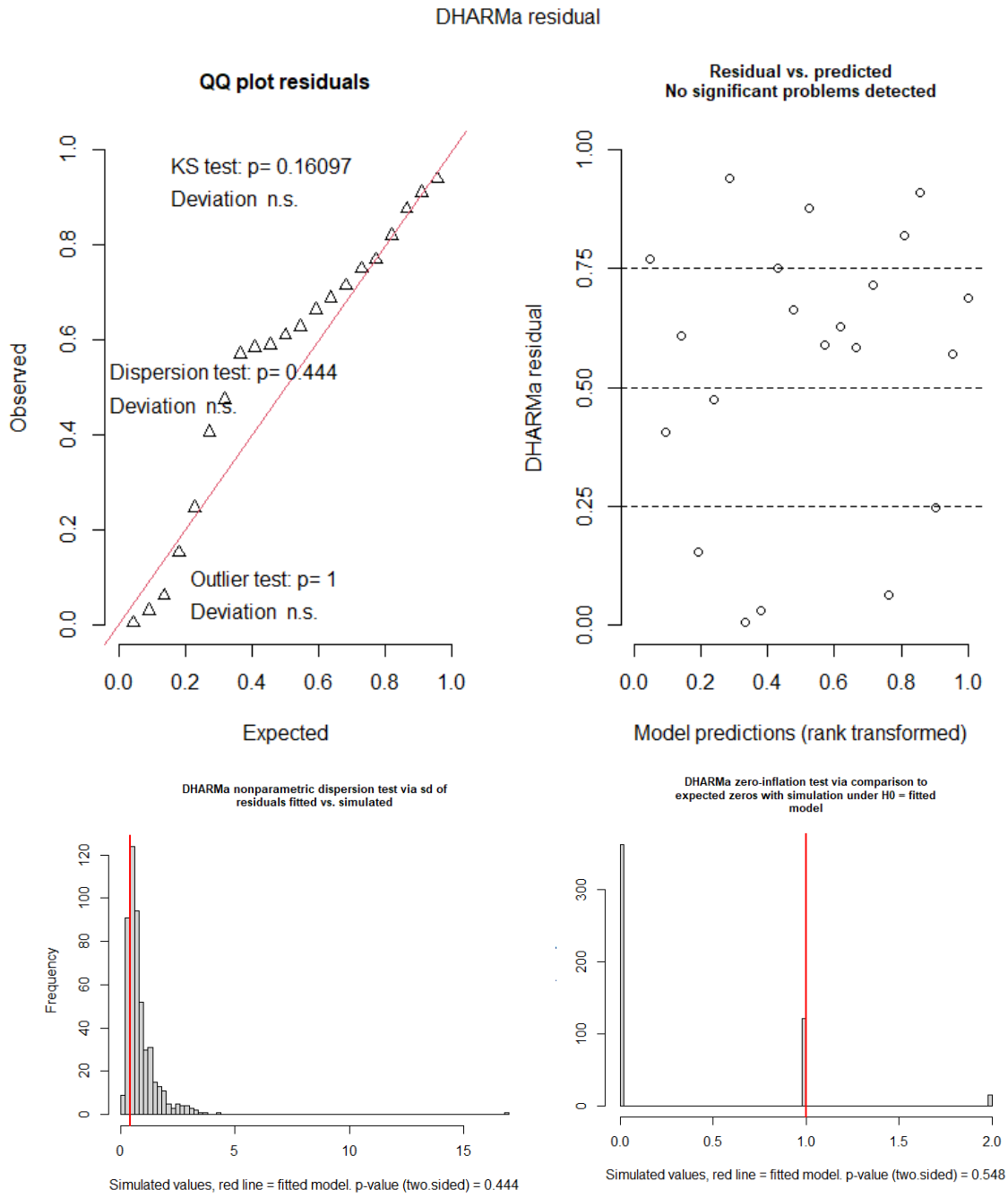


Figure 4.10. *Model validation plots for the NB2 GLM between pellet counts and the selected log-ratios of land covers, using DHaRMA package for R.*

The NB2 model with all selected LRs as predictors shown in Table 4.9 below indicated a significant effect of the LC3/LC5 predictor and explained 26% of the variation in pellet counts (marginal $R^2 = 0.258$). The incidence rate ratios of the predictors other than LC3/LC5 were not significantly different from 1, which corresponds to an absence of effect, due to large confidence intervals related to the small sample size and large variability (Table 4.9).

Table 4.9. Exponentiated NB2 GLM for pellet counts (y) and the linear combination of all selected land cover log-ratios (x). NB2 shape parameter = 0.968.

	pellets_counts		
Predictors	*Incidence Rate Ratios*	*CI*	*p*
(Intercept)	65.40	15.90 – 269.01	**<0.001**
LC1/LC3	0.68	0.32 – 1.44	0.316
LC3/LC5	0.49	0.26 – 0.94	**0.031**
LC4/LC5	0.92	0.49 – 1.74	0.809
LC2/LC5	1.60	0.59 – 4.35	0.354
Observations	21		
R^2 conditional / R^2 marginal	NA / 0.258		

The corresponding GLM equation is as follow:

$$log(y) = 4.18 - 0.38\log_2\left(\frac{LC1}{LC3}\right) - 0.71\log_2\left(\frac{LC3}{LC5}\right) - 0.08\log_2\left(\frac{LC4}{LC5}\right) + 0.47\log_2\left(\frac{LC2}{LC5}\right) \quad (10)$$

Equation 10 shows that the coefficient of $\log_2$(LC3/LC5) is the estimated effect on log(y) when doubling the ratio of LC3 over LC5, which corresponds to a one-unit increase in the log-ratio of base 2, while keeping other LRs constant. The acyclic connected graph (Figure 4.8) indicates that to keep LC1/LC3 constant, LC1 must increase by the same factor as LC3, while LC4 and LC2 have to decrease by the same factor as LC5 to keep LC4/LC5 and LC2/LC5 constant.

Hence, the effect size of the predictor LC3/LC5 indicated a significant effect of an increase in herbaceous and rock covers, along with a decrease in low-lying shrubs, refuge shrubs and bare soil covers, over pellet counts. The effect sizes of the other selected LRs correspond to the effect of a doubling of their respective numerator relatively to the rest of the land covers. This is because for each of these LRs, a one-unit increase while keeping other LRs constant implies that solely their respective numerator increases by a factor while the other land covers all decrease by a common factor, as shown by the acyclic graph in Figure 4.8. For example, a one-unit increase in $\log_2$(LC1/LC3) implies that the ratio of LC1 over LC3 doubles, so that LC1 increases by a given factor while LC3 decreases by another given factor. LC3 is connected to LC5, and LC3/LC5 must remain constant, hence LC5 has to decrease by the same factor than LC3. Similarly, LC2 and LC4 have to decrease so that the two remaining LRs are also kept constant, and hence, both have to decrease by the same common factor as LC5 and LC3. Therefore, a one-unit increase in $\log_2$(LC1/LC3) describes the effect on relative pellet density of doubling the ratio of rock relatively to all other land covers, for example, when passing from 20% in each land cover to 33% of rock and 16.5% of each other land covers at one site. This reasoning applies similarly for the two remaining ratios. In other words, the three non-significant effect sizes of the full model also yielded interesting information. They suggest that a relative increase of only bedrock, bare soil or refuge shrubs cover, respectively, did not significantly influence relative rabbit density at the observed sites.

The full model had a significantly higher AICc (250.7) than the null model (delta AICc = 9.1, Table 4.10), most likely due to the small sample size and large number of predictors. The full model is important to interpret the relationships between LRs in the context of regression as described, but due to the small sample size, it was decided to proceed with predictor selection in order to assess the effect of overfitting and determine if the predictor identified as significant

in the full model could keep some explanatory power in a reduced model. Thus, comparison of nested models with one, two and three predictors along with the null and full models were performed using AICc criterion (Table 3.10).

Table 4.10. Selection of the NB2 GLMs with lowest AICc and AICc weight > 0.05, for pellet counts (y) and combinations of one, two and three selected land cover log-ratios (x).

Model formula	K	AICc	Delta AICc	Model Likelihood	AICc weight	Log-Likelihood	Cumulative weight
pellets_counts ~ `LC3/LC5`	3	241.27	0.00	1.00	0.29	-116.93	0.2862801
pellets_counts ~ 1	2	241.67	0.40	0.82	0.23	-118.50	0.5204921
pellets_counts ~ `LC1/LC3`+`LC3/LC5`	4	244.08	2.82	0.24	0.07	-116.79	0.5905569
pellets_counts ~ `LC1/LC3`	3	244.22	2.95	0.23	0.07	-118.40	0.656053
pellets_counts ~ `LC3/LC5`+`LC4/LC5`	4	244.26	2.99	0.22	0.06	-116.88	0.7201143
pellets_counts ~ `LC3/LC5`+`LC2/LC5`	4	244.29	3.02	0.22	0.06	-116.89	0.7834742
pellets_counts ~ `LC4/LC5`	3	244.29	3.02	0.22	0.06	-118.44	0.8467014
pellets_counts ~ `LC2/LC5`	3	244.34	3.07	0.22	0.06	-118.46	0.9084973

The model using LC3/LC5 as the sole predictor had a marginally better AICc (delta AICc = 0.4) than the null model. It explained a similar share of variance in pellets to the full model (marginal R^2 = 0.236) but the predictor was not significant at the 0.05 level, because its upper confidence interval was slightly above one in the exponentiated expression of the model (Table 4.11), implying that its negative effect on pellets was not always present.

Table 4.11. Exponentiated NB2 GLM for pellet counts (y) and the retained land cover log-ratio (x). NB2 shape parameter = 0.923.

	pellets_counts		
Predictors	*Incidence Rate Ratios*	*CI*	*p*
(Intercept)	57.25	28.42 – 115.30	**<0.001**
LC3/LC5	0.65	0.41 – 1.02	0.059
Observations	21		
R^2 conditional / R^2 marginal	NA / 0.236		

The prediction plot of the single predictor model illustrates a non-linear decrease in the predicted pellet density as the proportion of herbaceous cover to low-shrubs cover (LC3/LC5) increased (Figure 4.11). The intercept indicates that when LC3/LC5 was null, which corresponds to an equal surface of herbaceous and low-shrubs covers, the expected pellet count was 57.25 pellets with a rather large 95% confidence interval ranging from 28.4 to 115.3 pellets per site. The incidence rate ratio suggests that for a one-unit increase of $\log_2$(LC3/LC5), which corresponds to doubling the ratio of herbaceous cover (LC3) to low-shrubs (LC5) cover, there was an average decrease of 35% in pellet counts.

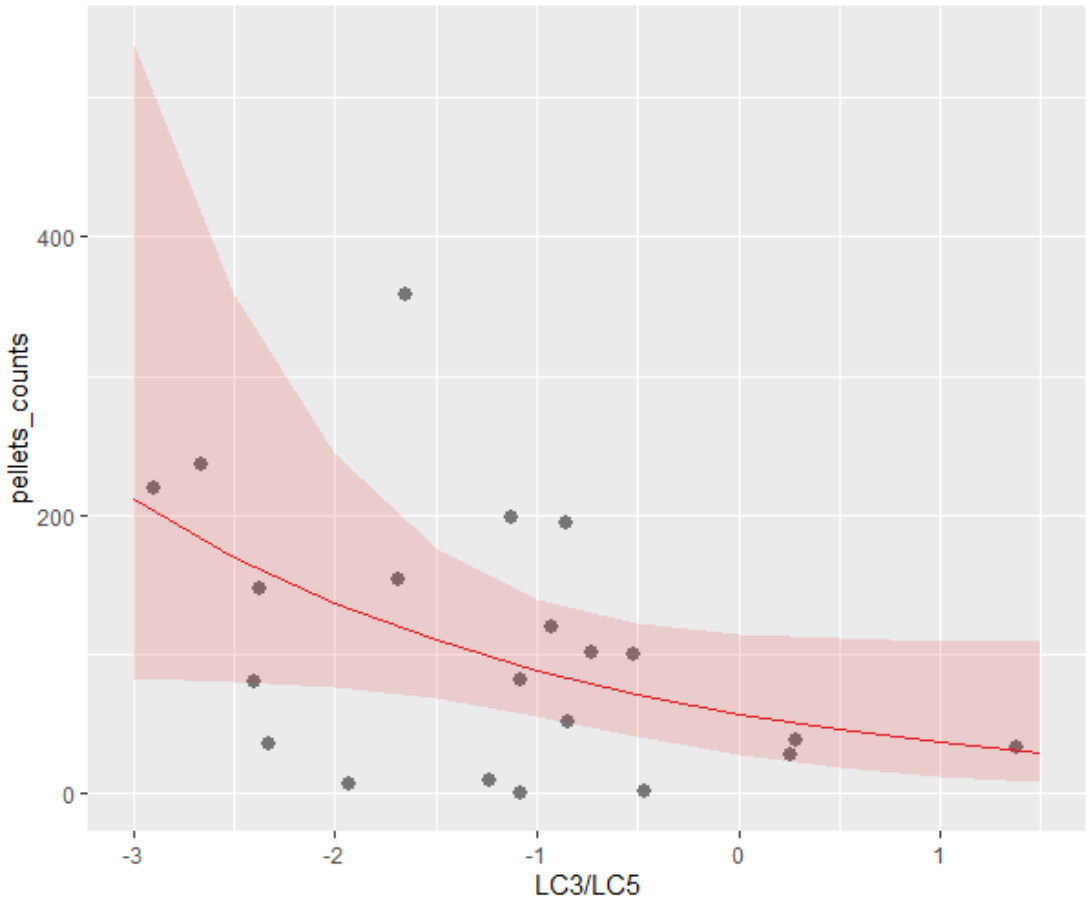


***Figure 4.11**. Predictive plot for the NB2 GLM with pellet counts (y) and the retained land cover log-ratio (x), which is in fact $\log_2$ (LC3/LC5). Red line is the predicted mean pellet count, red surface is the predicted confidence interval, and black dots are observed pellet counts.*

To verify further the role of LC3/LC5 in the land cover composition, a comparison of the predictions from the full model (i.e. including all selected LRs) and single predictor model (LC3/LC5) was carried out. Figure 4.12 below shows that the full model had large residuals, and notably could not predict null pellet counts, while a similar trend is observed for the single model albeit with larger residuals. The predictions of both models were moderately correlated (R^2=0.72, Figure 4.12), suggesting that the additional log-ratios of the full model somewhat contribute to the accuracy of the predictions. Both models had very similar residuals (R^2=0.90), indicating that the predictive errors of the full model are mostly explained by the single model, but the residuals of the latter were more randomly distributed, indicating less heteroscedasticity. Overall, these results suggest that the NB2 model with a single predictor (LC3/LC5) captured most of the influence of the land cover composition on pellet counts at the observed sites. These conclusions were further confirmed by a likelihood-ratio test between the full and single models that indicated an absence of difference in goodness-of-fit between the two models (Chi-square=1.1089 for 3 df, p-value=0.7749).

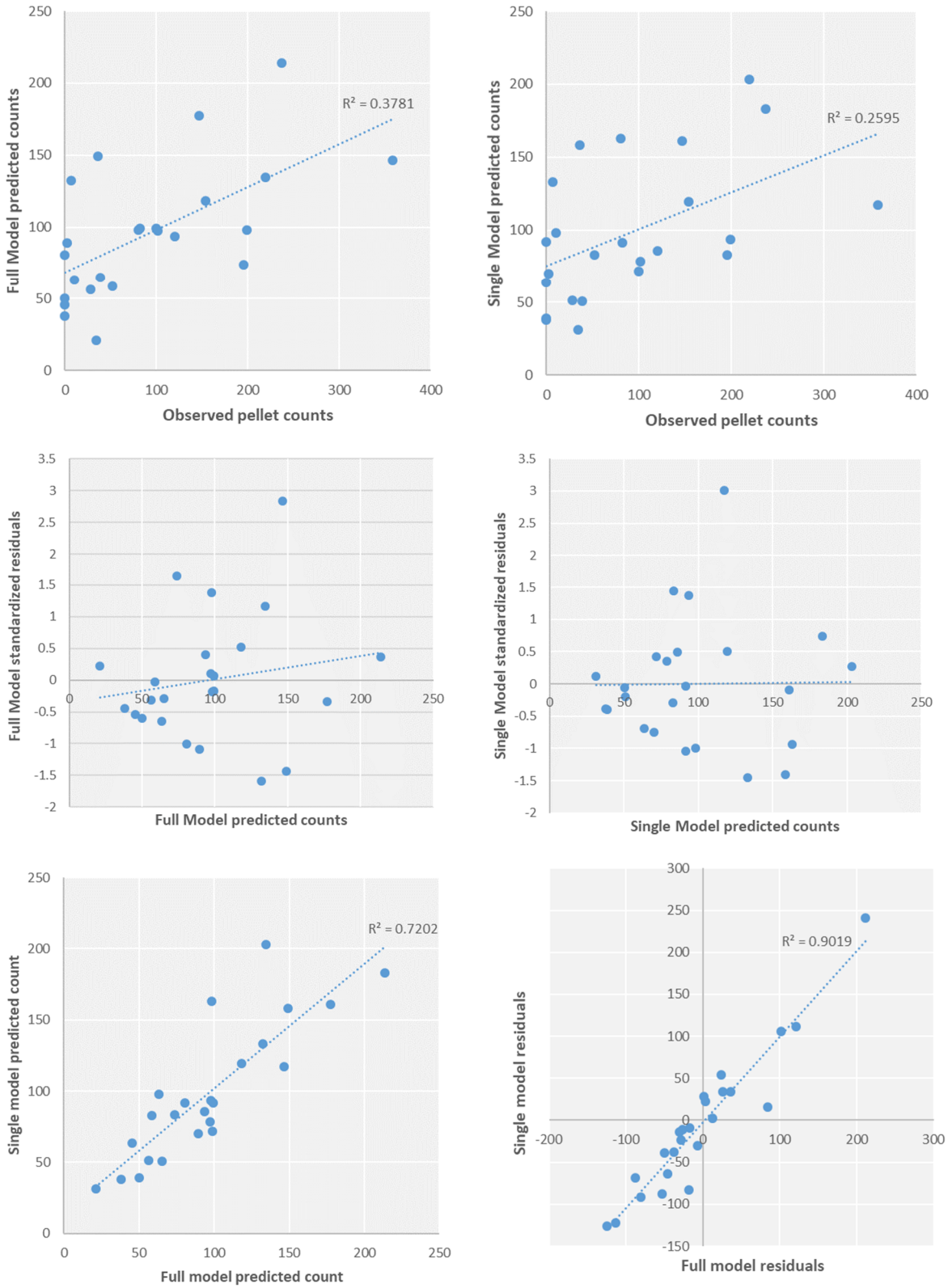


Figure 4.12. Linear regression scatterplots between the full and single NB2 GLM for pellet counts and land cover log-ratios. Top row: observed versus predicted pellet counts. Middle row: Predicted counts versus model residual errors. Bottom row, left: Predicted counts of full versus single models. Bottom row, right: Residuals of full versus single models.

4.4.3 Sensitivity analysis

The two land covers (LC3 and LC5) retained in the single model were those with the highest producer and user errors between each other (see section 4.2.3). Over the 1290 points used in the accuracy assessment, 12% of the points classified as low shrubs (LC5) belonged in fact to the herbaceous (LC3) class, while 4% of the points classified as herbaceous were in fact low shrubs. In this context, it was interesting to assess the extent to which the identified relationship between LC3/LC5 and pellet density could be due to classification errors.

Given the above, an exploration of the sensitivity of the findings to misclassification was conducted. To assess the robustness of the results against potential inaccuracies in land cover classification, the original land cover dataset was corrected by redistributing the surface areas of each land cover type based on the error rates presented in Table 4.12, themselves derived from the global accuracy assessment detailed in Table 4.4b. As this global approach did not take into account the variability in misclassifications between sites, it cannot be considered as a full sensitivity analysis but only as an indicative test. This correction led to a reduction in surfaces of LC5 (low shrubs) whereas those of LC3 (herbaceous cover) were increased, bringing the values of these two classes closer to each other, ultimately decreasing the variability in the LC3/LC5 ratio across sites. Such correction was expected to result in a narrower range of values for this predictor and thus, a potential reduction of its explanatory power in the GLM.

Table 4.12. Error rate matrix for land covers classification. Rows indicate the proportion of pixels removed from a given class. Columns indicate the proportion of pixels added to a given class. Example: 93.1% of original LC1 pixels are retained in the corrected LC1 class, while 0.99% are added to the corrected LC2 class, and 5.26% of the original LC2 pixels are added to the corrected LC1 class.

		Corrected				
		LC1	**LC2**	**LC3**	**LC4**	**LC5**
Original	LC1	93.10%	0.99%	0.00%	1.97%	3.94%
	LC2	5.26%	91.81%	2.92%	0.00%	0.00%
	LC3	0.38%	1.13%	93.96%	1.13%	3.40%
	LC4	4.32%	0.00%	5.76%	89.21%	0.72%
	LC5	6.70%	3.24%	11.88%	0.86%	77.32%

However, when the GLM procedure was performed again on the corrected land cover dataset, the model with LC3/LC5 as the sole predictor was still retained as the only model with a slightly lower AICc (delta = 0.5) than the null model. Surprisingly, the model with corrected data had a slightly higher R^2 value (marginal R^2 = 0.24) than with the original dataset, while the predictor had a lower incidence rate ratio and confidence interval, leading to a marginal increase in significance as indicated by the p-value that shifted just below the indicative 0.05 threshold (Table 4.13).

Table 4.13. Exponentiated NB2 GLM for pellet counts (y) and the retained land cover log-ratio (x), using the dataset corrected for misclassification errors. NB2 shape parameter = 0.923.

	pellets_counts		
Predictors	*Incidence Rate Ratios*	*CI*	*p*
(Intercept)	71.76	42.15 – 122.14	**<0.001**
LC3/LC5	0.55	0.30 – 1.00	**0.048**
Observations	21		
R^2 conditional / R^2 marginal	NA / 0.240		

The GLM analysis conducted in this section indicated a marginally significant (p-value < 0.1) negative association between the proportion of herbaceous to low-lying shrubs covers and relative rabbit density in the Maltese garrigue, suggesting that the null hypothesis of the second research question of this dissertation could be rejected with weak evidence. However, the large unexplained variance in the above model suggested that there were unaccounted factors affecting pellet density. The next section answers the third research question, which goes beyond the natural environment considered until now, in order to test if the unexplained variation of pellet density could come from anthropogenic sources.

4.5 Relative rabbit density and anthropogenic features

4.5.1 Descriptive statistics

The three variables of interests contained surfaces measured in metres squared (m^2), and only the surface of roads had zero values found at every scale (Table 4.14).

Table 4.14. Descriptive statistics for the measured anthropogenic variables at each scale.

Variable	Mean	St.Dev.	Min	Max	Median	Zeros count
roads_100	511	710	0	2088	61	9
roads_200	2573	2755	0	9621	2338	4
roads_500	14879	10910	0	41277	13729	1
paths_100	1987	1037	314	4359	1989	0
paths_200	5129	1842	2016	8197	5082	0
paths_500	22364	7945	11648	41931	22597	0
rubbles_100	1540	1055	35	3907	1458	0
rubbles_200	5558	2901	924	10487	5006	0
rubbles_500	28402	11727	8824	53694	30053	0

4.5.2 Model

An initial Poisson GLM was fitted using the three standardized predictors at their largest scale (500m buffer). Significant overdispersion, zero-inflation and departure from normality of the residuals were detected with similar results to the previous Poisson GLMs presented above. The mean-variance binned plot for the Poisson model shown in Figure 4.13 below suggested a quadratic trend (NB2) as the best fit to the data.

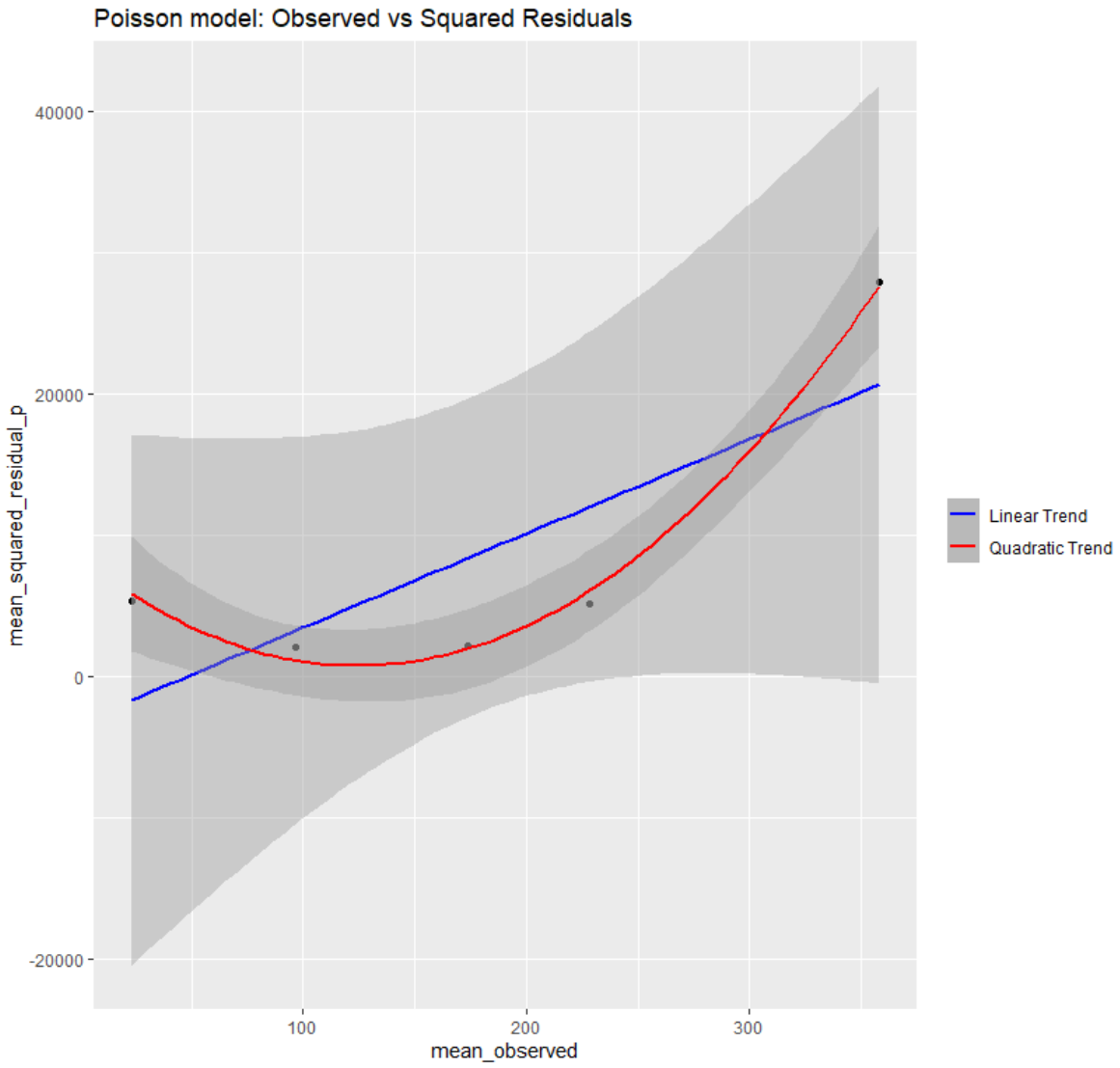


Figure 4.13. *Mean observed pellet count (x) against mean squared residuals (y) of the Poisson GLM between pellet counts (binned by equal group width) and the three anthropogenic variables at their largest scale as explanatory variables. The blue line indicates a regression using the NB1 formula, while the red line indicates a regression using the NB2 formula.*

Hence, a NB2 model was fitted with the three 500m predictors and n=500 simulated samples. Residuals of the model did not show any significant issue of departure from normality, dispersion nor outliers (Figure 4.14). Expected zeros from the fitted model were higher than the observed zeros without being significant (ratio=3.36, p-value=0.53), while the dedicated dispersion test reported insignificant underdispersion of the model (0.42, p-value= 0.37).

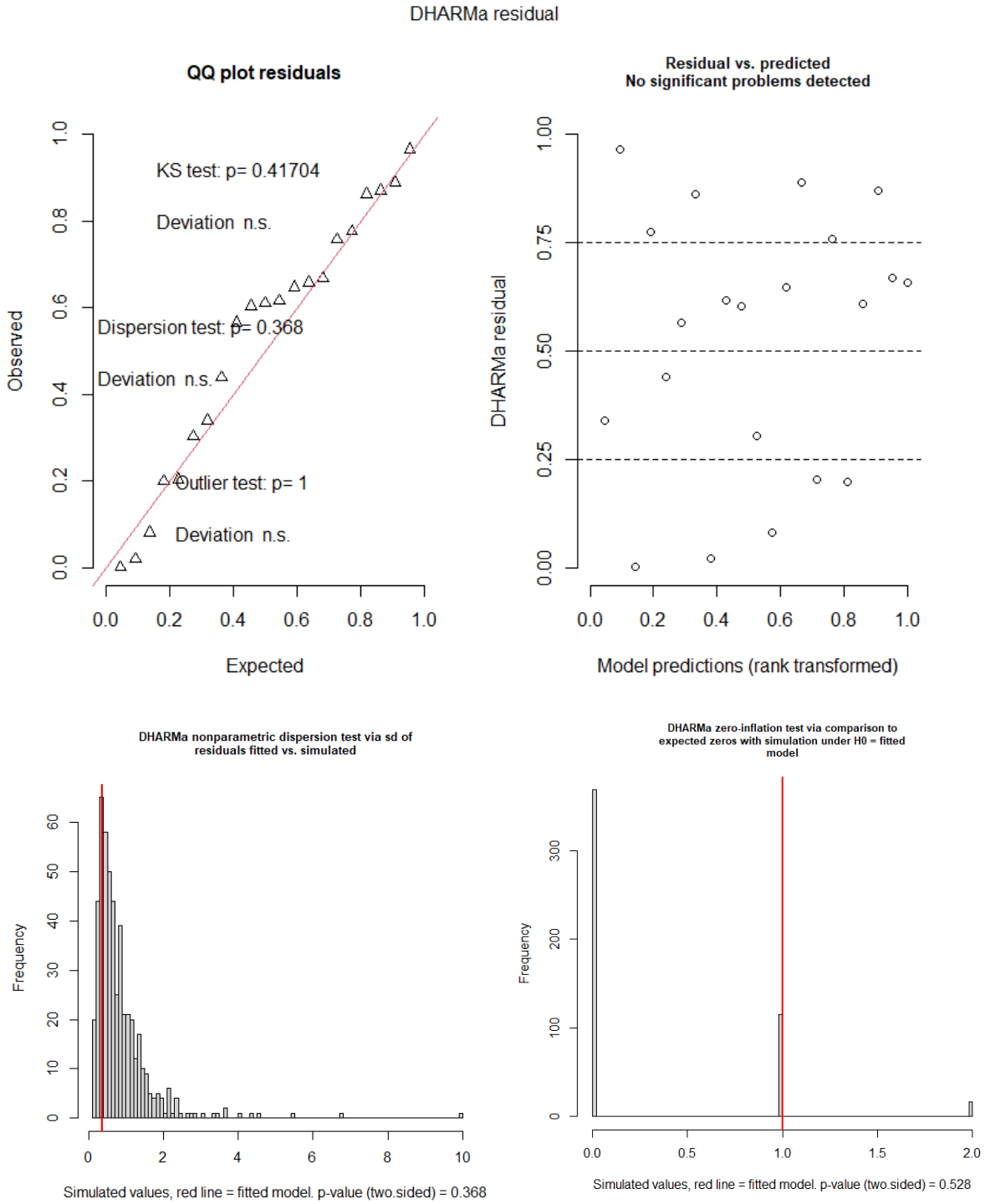


Figure 4.14. *Model validation plots for the NB2 GLM between pellet counts and the selected log-ratios of land covers, using DHaRMA package for R.*

Models for all combinations of the three predictor categories and their three different scales were compared using AICc criterion, including the null model. Table 4.15 below presents the models with AICc weight > 0.05.

Table 4.15. Selection of the NB2 GLMs with lowest AICc and AICc weight > 0.05, for pellet counts (y) and combinations of one, two and three anthropogenic variables (x) as predictors.

Model formula	K	AICc	Delta AICc	Model Likelihood	AICc weight	Log-Likelihood	Cumulative weight
pellets_counts ~ roads_200 + paths_500	4	236.40	0.00	1.00	0.19	-112.95	0.19
pellets_counts ~ roads_200 + paths_500 + rubbles_500	5	237.52	1.12	0.57	0.11	-111.76	0.30
pellets_counts ~ roads_200 + paths_500 + rubbles_200	5	237.95	1.55	0.46	0.09	-111.98	0.38
pellets_counts ~ roads_200 + paths_500 + rubbles_100	5	238.36	1.96	0.38	0.07	-112.18	0.45

The first model of Table 4.15 above was preferred for its lower number of predictors in order to avoid overfitting. It had an AICc difference of 5.3 with the null model and the two predictors were found to have similar negative effects on pellet counts, while explaining 48% of the variation in pellet counts (marginal R^2 = 0.48, Table 4.16).

Table 4.16. Exponentiated NB2 GLM for pellet counts (y) and the extent of impermeable roads and permeable paths (x) as predictors. NB2 shape parameter = 1.22.

	pellets_counts		
Predictors	*Incidence Rate Ratios*	*CI*	*p*
(Intercept)	82.93	56.15 – 122.48	**<0.001**
roads 200	0.52	0.32 – 0.86	**0.011**
paths 500	0.51	0.34 – 0.78	**0.002**
Observations	21		
R^2 conditional / R^2 marginal	NA / 0.481		

The prediction plots for the anthropogenic model show that as the surface of roads and paths increased, the count of pellets per site tended to decrease non-linearly (Figure 4.15). The intercept indicates that in the absence of roads and paths, under 200m and 500m around sites respectively, the expected pellet count was 82.9 pellets on average with a 95% confidence interval ranging from 56.2 to 122.5 pellets per site. An increase of one standard deviation in the surface of impermeable roads found within and around sites up to 200m, corresponding to an increase of around 2750m², decreased the count of pellets per site by 48% (Table 4.16). Similarly, an increase of one standard deviation in the surface of permeable paths found within and around sites up to 500m, corresponding to an increase of around 7940m², decreased the count of pellets per site by 49%.

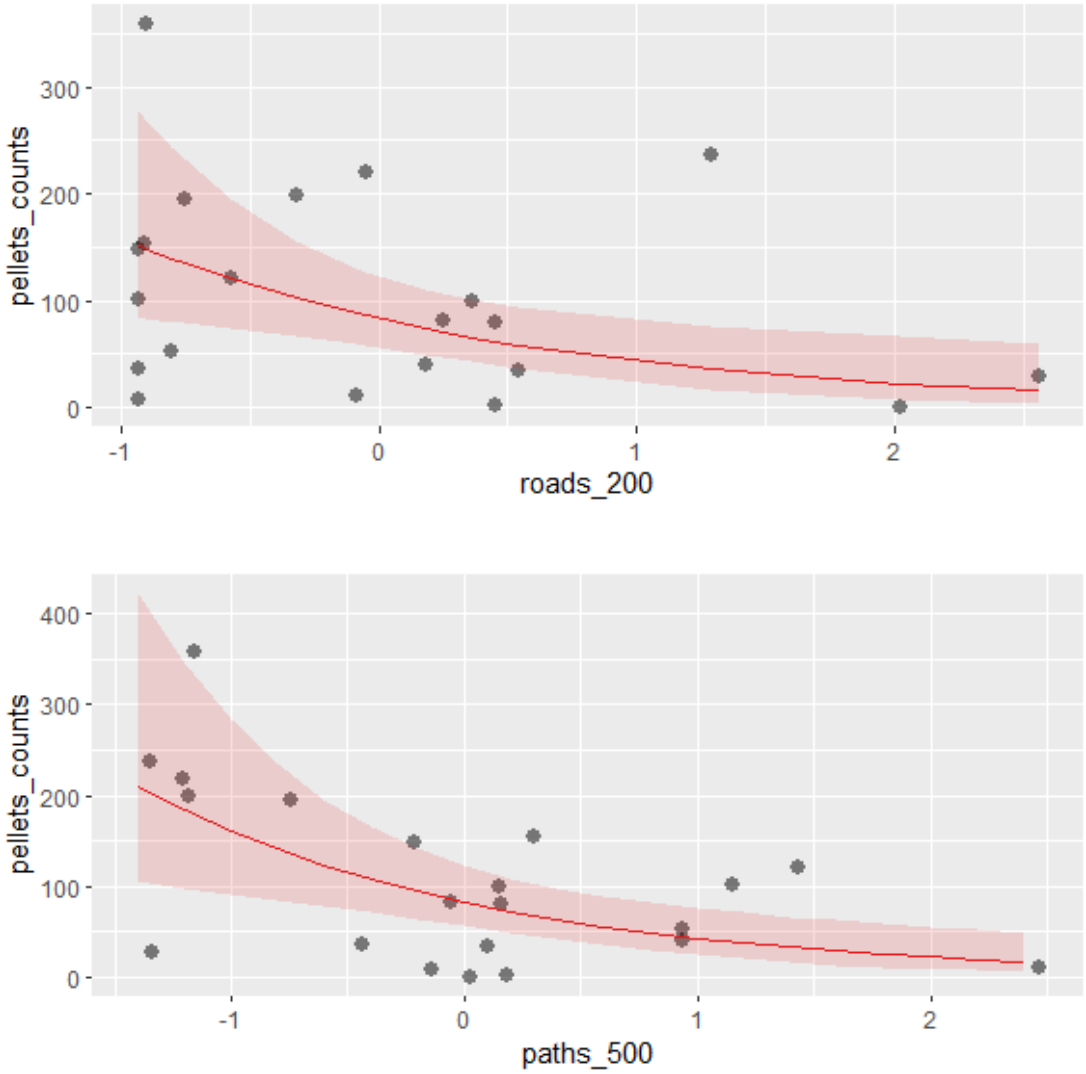


Figure 4.15. *Predictive plots from the retained NB2 GLM with pellet counts (y) and (top) the extent of roads within 200m from sites (x), and (bottom) the extent of paths within 500m from sites (x). Solid red line is the predicted mean pellet count, red surface is the predicted confidence interval, and black dots are observed pellet counts. The x axis displays standardized values.*

The GLM analysis conducted in this last section indicated moderate (p-value < 0.05) to strong (p-value < 0.005) evidence of an influence of the measured anthropogenic features on rabbit relative density in the Maltese garrigue, suggesting that the null hypothesis of the third research question of this dissertation can be rejected with moderate evidence.

4.6 Results summary

In this chapter, the accuracy assessment of the image classification process was reported as satisfactory (Kappa = 0.83). A separate GLM procedure was carried out for each research question and stopped after the predictor selection phase, this is because the models were used for inference rather than prediction. The null hypothesis of an absence of relationship between shrubs and relative rabbit density could not be rejected. The null hypotheses for the second and third research questions were rejected due to the presence of significant predictors in the models, indicating probable effects of land covers and anthropogenic features on relative rabbit density.

5. Discussion

The European rabbit represents a conservation paradox: it is crucial for ecosystem and predator support in its native European habitats while it is considered as an invasive pest subjected to rigorous control measures in introduced ranges such as Australia and New Zealand, with historical eradication campaigns carried out in its native range as well (Cooke 2008; Delibes-Mateos et al. 2011; Lees & Bell 2008). This paradox is also present in Malta, where the wild European rabbit is traditionally viewed as an agricultural pest and occasionally labelled as invasive, receiving protection only under regulations focused on game management, despite its central role in national identity (Cassar 1994). Moreover, its ecological role within Malta's Mediterranean semi-arid ecosystems, which share some characteristics with the native range of the species, has never been studied. To this end, the present study carried out a survey of 24 sites found in seven open garrigue areas on the main island of Malta over a single season, to explore selected ecological relationships between rabbit density and selected environmental factors known to be critical in its native range. The findings did not reveal any significant association between the shrub vegetation and rabbit density but seemed to point instead to a weak association with the land cover composition and a moderate

association with anthropogenic features. The methods used and the small sample size studied here call for caution in interpreting the statistical results.

5.1 Shrubs as refuge

Positive associations between shrub cover and rabbit density have been widely reported in the native range of the European rabbit (Beja et al. 2007; Calvete et al. 2004; Carvalho & Gomes 2004; Fernandez 2005; Monzon et al. 2004). In these studies, shrubs are interpreted as refuges from predation allowing rabbits to hide temporarily while grazing, with an accrued importance in landscapes where burrowing is more difficult due to soil's characteristics (Carvalho & Gomes 2004). The latter authors showed that rabbit abundance was associated with gaps in rocks and shrub cover in areas where warren construction was hindered due to "squeletic soils scattered with rocks of various sizes" (p. 71), a quote particularly adapted to describe the soils of the Maltese garrigue. Hence, the importance of shrubs for refuge availability in Malta was associated with the alternative hypothesis of the first research question, given the local edaphic conditions.

Conversely, the predator-prey dynamics in Malta present a contrasting scenario to that observed in the rabbit's native range, with dozens of predators depending on the rabbit as a food source in Spain (Delibes-Mateos et al. 2007). In Malta, rabbit predation has not been quantified yet but is considered to be low, with sparser human hunting and seasonal aerial predation from migrating raptors compared to the Iberian Peninsula. This reduction in predation pressure might minimize the role of refuge availability for the rabbit distribution in Malta and was therefore associated with the null hypothesis of the first research question.

Parameters such as vertical structure (Beja et al. 2007), interspersion with grass patches and total canopy cover at home-range scale (Fernandez 2005) have been identified as main drivers

of the shrubs' refuge function. Therefore, it was quantified here through the number of individual shrubs, their degree of aggregation and the total shrub cover per hectare, chosen as proxies for canopy cover and interspersion, while individual shrub height and area were selected as proxies for shrub structure. These parameters were also selected for their feasibility according to the resource limitations of this study.

However, none of these variables appeared to have a noticeable effect on relative rabbit density in the studied sites of the present dissertation, preventing the rejection of the null hypothesis of the first research question. This outcome suggests that shrub cover as a whole might not be as important in the Maltese garrigue than in its native Mediterranean semi-arid ecosystems for the wild European rabbit distribution, possibly due to the lack of significant predation pressure.

The current study could not distinguish between different species of shrubs, and this might well be a confounding factor in the reported results. While some previously mentioned studies did report an effect of an undifferentiated shrub cover, Beja et al. (2007) reported contrasting effects of different shrub species based on their differing vertical structure. Shrubs with a sparse ground-level vegetation and dense overhead canopy (i.e., *Cistus ladanifer*) were positively associated with rabbit density, while shrubs with a denser ground vegetation (ericoids) had a negative association with rabbit density (Figure 5.1).

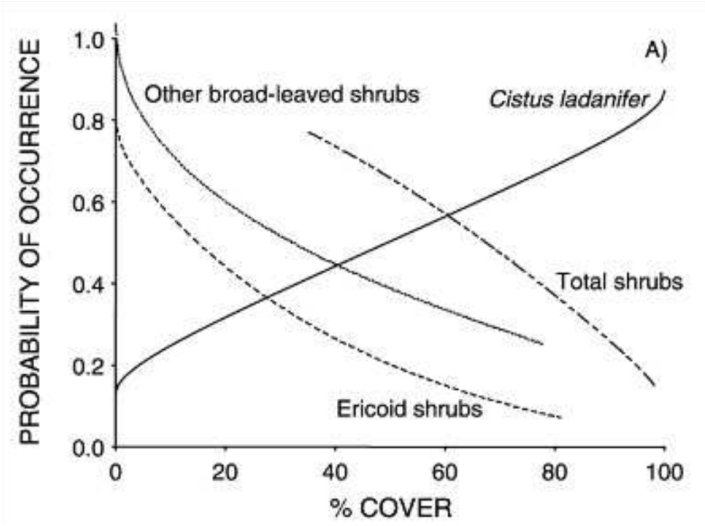


Figure 5.1 *Logistic regression models for the effects of variables characterising the scrub layer on the probability of occurrence of rabbits in Mediterranean scrublands in SW Portugal, 1997. From Beja et al. (2007, p. 33).*

Various families of shrubs co-occur in the assemblages of Maltese garrigue, notably the ericoid *Erica multiflora*, while the Cistaceae family has two recorded species in Malta (*C. creticus* and *C. monspeliensis*), and others notable species occur such as *Pistacia lentiscus* and the Euphorbiaceae members (Brullo et al. 2020). The lack of distinction between shrub species might have masked contrasting effects of different families' covers over relative rabbit density, leading to the absence of effects of the overall shrub structure reported here, whether it be the number or the total surface of shrubs. Significant differences between shrub species in terms of internal structure (e.g. spines, stems density), chemical composition (e.g. effects on soil), or feeding potential (e.g. toxicity/palatability) could possibly influence the effect of shrubs on rabbit spatial distribution and shape some of their habitat preferences in Malta.

It might be due to the higher availability of other types of refuge such as rock cracks as observed by Carvalho and Gomes (2004), which may be offered locally by rubble walls, ruins and geomorphological features. Rabbits are known to select favourable above-ground

landscape configurations offering a mixture of shrubs and pasture in shallow-soil Mediterranean ecosystems at the expense of warren building potential (Serronha 2014). In Malta, rubble walls not only offer convenient refuge in interstices, but also generate deeper tracts of soil along their uphill edges due to agricultural terracing (Sultana 2015). This practice has largely shaped the Maltese landscape and its effect is still observable even where terraces are abandoned. The presence of deeper soil in the garrigue may counter-balance the importance of shrubs for refuge availability by facilitating localized warren construction in landscapes where the overall soil depth should not allow it.

5.2 Land covers composition

Land covers have been widely reported as influential for the European rabbit distribution in native Mediterranean environments (Arques et al. 2014). In particular, the extent of herbaceous cover is associated with food availability, the extent of hard rock and bare soil covers are associated with warren suitability, and canopy cover is associated with refuge availability and predation avoidance (Carvalho & Gomes 2004; Delibes-Mateos et al. 2010; Kontsiotis et al. 2013; Lombardi et al. 2003; Martins et al. 2003; Perez et al. 2008).

This study found weak evidence of an influence of vegetation and landscape composition over relative rabbit density at the observed sites in the Maltese garrigue, but the effect was found to be robust to misclassification errors. The variation in pellet density across sites was partially explained (marginal R^2=0.26) by the dynamic between the herbaceous and rock covers contrasting with low-lying shrubs, refuge shrubs and bare soil covers. In fact, the effect of land covers was mostly attributable to the ratio of herbaceous to low-lying shrubs covers (marginal R^2=0.24), with a negative effect of the herbaceous cover and a positive effect of the low-lying

shrubs cover However, there was only a small difference between the null and full model (delta AICc = 0.4), which suggests that the additional variables included in the full model may not significantly improve the model's ability to explain the variation in rabbit density across the observed sites. This indicates that while land cover types do play a role in determining rabbit density, other unmeasured environmental or biological factors might also be influential.

A hump-shaped relationship between herbaceous cover and rabbit density has been described in the literature because high herbaceous cover indicates more open areas with increased predation risks if interspersion with shrubs is insufficient, while low herbaceous cover is associated with limited quality-food supply (Delibes-Mateos et al. 2010; Lombardi et al. 2003). Here, a negative relationship between herbaceous cover (LC3) and relative rabbit density was observed, which might indicate that rabbits avoid areas with larger herbaceous vegetation cover even though it may provide ampler food supply. As indicated by the full model, herbaceous cover tended to increase along with rock cover in opposition to the three other classes. As detailed in section 2.1.3.1, the Maltese garrigue is often mixed with steppic patches/ecotones, which generally contain more grass species, less shrub species, more exposed bedrock and shallower soil. Hence, the uncovered relationship may suggest a negative effect of steppic-like patches of the garrigue on rabbit pellets density during the summer season. On the other hand, the low-lying shrubs (LC5) were positively associated with relative rabbit density along with refuge shrubs and bare soil. The refuge shrubs and soil covers are often considered as indicators of refuge availability (Delibes-Mateos et al. 2009), however, the low-lying shrubs considered in this study are unlikely to act as refuges due to their dense inner structure and maximum height similar to the rabbit (<25cm). Instead, low-lying shrubs cover may well have acted as an indicator of food availability during summer season. This land cover class was mainly separated from the refuge shrubs through differences in relative

elevation, while it was distinguished from the herbaceous cover mostly through spectral differences, because both classes occupied the same vertical strata. The low-shrubs land cover had a greyer and bluer tint than the herbaceous cover, as can be seen in Figure 4.2, which correspond to the overwhelming presence of *Thymbra capitata* in this class.

Wild *Oryctolagus cuniculus* is known to feed on less palatable vegetation depending on season and local primary productivity patterns (Martins, Milne & Frego 2002). Kontsiotis et al. (2015) reported an increase in browsing of *Thymbra capitata* by the rabbit in the phrygana during July, a behaviour also reported by Matrai, Altbacker and Hahn (1998) for *Thymus glabrescens* during summer in a Hungarian juniper forest. Consumption of *Thyme* during a controlled experiment in the same ecological context was shown to reduce litter size, increase pup mortality and induce mild intoxication symptoms in pups compared to control diets, but also revealed that this feeding behaviour was nonetheless transmitted from mother to pups (Altbacker, Hudson & Bilko 1995). On the other hand, Benlemlih et al. (2014) reported a mild antimicrobial effect of a diet with 3% supplemented *Thymbra capitata* for the European rabbit, which reduced mortality. This is corroborated by Ezzat Ahmed et al. (2020, p.561) who found that *Thyme* significantly improved "productive and reproductive performances" of the rabbit under hot climatic conditions at certain dosage. Hence, *Thyme* may be considered as an important part of the diet of the wild European rabbit in Mediterranean scrubland ecosystems during summer regardless of its mild toxicity, due to a more constant availability of younger plant parts (Martins, Malero & Nogales 2003), its potential benefits at certain concentrations in the diet and the transmission of feeding behaviours between generations.

The present research could reflect the observations reported above for the Maltese garrigue. The vegetation survey was carried out in February when primary productivity was at its peak due to the wet season, but the rabbit pellet survey was carried out in June when the herbaceous

cover had already started to dry up. However, *Thymbra capitata* is an evergreen plant with leaves available all year-round, a flowering phase from May to June while the seeds are mature one month after the bloom, thus in July and August at the peak of Maltese dry season (Rey & Saez 2002). Thus, at the time of the pellet survey (June 2023), thyme must have been the main available food source for the European rabbit in the Maltese garrigue, possibly explaining why more pellets were found at sites with higher shrubs to grass ratios.

Rabbit pellets can endure several months in place and can be considered as a proxy for rabbit density; however, it is not an absolute density indicator and should still be interpreted in the context of random pellet deposition during feeding (see section 2.2.3). Moreover, the Maltese wet season (i.e. winter) has severe storm events that can still occur in late spring, thus, pellet displacement and increased decay rates may have occurred between the vegetation and the pellet surveys. Therefore, it is safer to assume that the dung pellet survey indicated spatial feeding preferences related to the onset of the dry season in Malta. Sites with a high herbaceous cover in February had a larger dry vegetation extent in June, while the low-lying shrubs cover (comprised of mainly *T. capitata*) must have remained rather stable between the two surveys. Unfortunately, without seasonal observations of pellet density, it was not possible to determine if wild rabbits in Malta favour sites with greater low-lying shrubs cover throughout the year due to a more stable food availability, or if this selection occurs in summer due to a temporary shift in feeding patterns. The size of the observed sites is smaller than the average home range of *Oryctolagus cuniculus*, and some areas comprised sites potentially close enough to each other to be used by the same colony. Hence, it is possible that rabbits shifted their feeding locations to sites with more low-lying shrubs instead of those with more grass cover, primarily due to seasonal patterns of food availability in the Maltese garrigue.

T. capitata was found to be negatively correlated to plant species diversity and richness in the Maltese garrigue by Galea (2017), without accounting for therophytes. This author pointed to a 'shrub encroachment' effect of low, sprawling shrubs like thyme over the herbaceous vegetation. Moreover, she notes that "the same factors that cause an increase in the abundance of *Thymbra capitata* [may] also cause a decrease in the diversity of all other species", stressing the need for further experimental work to investigate causality (Galea 2017, p.80). The study also reported an increased influence of *Thymbra capitata* at the largest considered scale (125x125m), which correspond to the scale of the vegetation survey carried out here (100x100m). In that study, *Thymbra capitata* accounted for 27% of the variation in the plant communities that was not explained by the measured site characteristics. Thyme was hypothesized to "influence chemical and physical aspects of the environment that would be relevant to plant growth and […] to depress the growth rate of other plants in its vicinity" (Galea 2017, p. 82). But what may in turn drive the variations in cover of *Thymbra capitata* inside the Maltese garrigue?

In the present study, the LC3/LC5 ratio explained 24% of the variations in rabbit pellet density across sites, with more pellets found at sites with greater low shrubs cover. While no quantification of the spatial effects of rabbit grazing on thyme cover was ever done in Malta, it seems safe to assume that the cover of this species is less influenced by browsing/grazing than the herbaceous cover. Rabbits mostly eat new, growing stems and flowers from shrubs like thyme, while herbaceous species are usually entirely eaten, sometime including their roots (Delibes-Mateos et al. 2020). The vegetative phase of annual plants in Malta is rather short and concentrated around spring, while thyme grows continuously from spring across the summer (Brullo et al. 2020).

Hence, the influence of *T. capitata* on plant community composition reported by Galea (2017) could be at least partially driven by the effects of wild European rabbit's grazing. Sustained grazing can reduce herbaceous layer's regeneration over time through seed bank depletion and plant competition mechanisms (Pol, Sagario & Marone 2014). This effect was already suspected in a preliminary study of the present author that observed a disproportionate regeneration of the herbaceous layer in the absence of rabbit grazing on the island of Comino, where high rabbit densities occur (Tranchant 2022). Interactions between herbaceous and shrubby vegetation in Mediterranean ecosystems in the presence of grazing are complex and dynamic (Maestre et al. 2003; Madrigal et al. 2008; Osem et al. 2007). Shrubs offer protection from grazing by large species to the herbaceous plants in their immediate vicinity, however, they are unlikely to offer such protection from the rabbit due to its small size. In addition, the effect of grazing on the herbaceous layer could be worsened if rabbits cannot easily colonize new areas where the herbaceous vegetation has not yet been grazed. Rabbits generally do not disperse far from their native warren and are pushed to do so mainly by warren population densities and inter-male reproductive competition (Mykytowycz & Gambale 1965; Sneddon 1991). Therefore, spatially constrained rabbit population are likely to have time to alter the vegetation community before being pushed to colonize new areas. The reduction in herbaceous vegetation caused by rabbit grazing can in turn favour thyme and other sprawling shrubs in their competition for space in the karstic garrigue of Malta where available soil patches are rare, ultimately facilitating the shrub encroachment process suspected by Galea (2017).

Some important aspects of the rabbit ecology were not considered in this study, such as are the degrees of fragmentation of the landscape and interspersion of vegetation patch types. Virgos et al. (2003) found that rabbit abundance was correlated to shrub cover but not to grass

cover in fragmented areas, whereas the proportion of both covers was influential in continuous (i.e. less fragmented) habitats. They also found that macrohabitat factors had more explanatory power than microhabitat factors on the rabbit's spatial distribution modelling. In the present study, only the degree of shrub clustering was accounted for, but there could be a mosaicking effect at play in the Maltese garrigue where large continuous patches of homogeneous vegetation could have a different effect than fragmented patches on rabbit density, modulated by patch types.

In summary, the findings of this research related to the second research question seem to concur with the observations of habitat preferences of the wild European rabbit in its native range. In the Maltese garrigue, the wild rabbit appears to favour sites offering some type of balance between year-round food availability and some refuge availability, albeit with a reversed importance of these two factors compared to the Iberian Peninsula. It also seemed to avoid more steppic sites with more open landscapes that offer less refuge potential and more varying food availability. Moreover, the results highlighted the potential effect of rabbit's grazing on plant community dynamics in the Maltese garrigue, where it is suspected to facilitate shrub encroachment of *Thymbra capitata* at the expense of the herbaceous layer.

5.3 Anthropogenic influence

The third research question of this thesis was carried out as a complementary step to the study of natural factors of the two other research questions. Given the degree of artificialization of the Maltese archipelago, anthropogenic features were suspected to play a larger role in the spatial distribution of the wild rabbit than in its native range. The surface of impervious roads was taken as an indicator of connectivity barriers, the surface of permeable

paths was taken as an indicator of disturbance from human recreation, and the surface of rubble walls was taken as an indicator of refuge availability beside vegetation.

Human infrastructure has a detrimental effect on biodiversity mostly due to the habitat fragmentation and land use change it induces (Benítez-López, Alkemade & Verweij 2010; Coffin 2007; Maxwell et al. 2016). Transport infrastructure in particular poses several threats to biodiversity by creating connectivity barriers that can even affect gene flow (Forman & Alexander 1998; Gerlach & Musolf 2000). It has contrasted effects on mammals depending on its spatial configuration, with increased roadkill along minor roads (Van Langevelde, Van Dooremalen & Catharinus 2009), traffic intensity negatively associated with predator avoidance (Bautista et al. 2004), or effects on specific predator-prey relationships such as increased predator roadkill due to increased prey abundance (Silva et al. 2019). In fact, railways and roads have generally been found to contribute positively to European rabbit abundance because they offer suitable soil for warren construction along verges (Rouco et al. 2019) and reduce predation risks due to the side-effects of predator roadside avoidance and roadkill (Barrientos & Bolonio 2009; Bautista et al. 2004; Planillo & Malo 2013). Similar contrasting effects have been noted for the influence of human recreation on wildlife in open areas. On one hand, trail frequentation by humans and their pets have led to shifts in daily activity patterns, occupancy and habitat use for wildlife (Lewis et al. 2021). On the other hand, these changes in predator behaviour can sometimes be favourable to prey species, such as lagomorphs, which can use areas of reduced predator activity to their benefit (Caldwell & Klip 2022). Lastly, rubbles walls in Malta have been traditionally used as a means to prevent soil erosion and loss (Sultana 2015). As such, they often generate an accumulation of soil on their uphill side which could be used by rabbits for warren construction in areas where soil is shallow

otherwise. In addition, rubble walls often present cracks and cavities which rabbits may use to find refuge, even though it is unlikely to be used as a replacement for warrens.

In this study, the extent of permeable paths (i.e. trails) and impermeable roads (i.e. motorways) were found to be negatively associated with relative rabbit density in the Maltese garrigue, explaining together 48% of the variability in pellet counts with the negative binomial GLM. The extent of rubble walls showed a mild positive effect, albeit not significant enough to be retained in the final model.

Konstsiotis et al. (2013) reported a similar negative effect of paved roads on rabbit densities on the island of Lemnos, Greece, which shares some of the environmental and anthropogenic characteristics of Malta. Fernandes (2012) found a negative association between some anthropogenic features, including roads and built surfaces, and the number of rabbit warrens in Portugal. Similarly, Planillo and Malo (2013) found lower rabbit densities closer to motorways in Spain. The reduction of rabbit density closer to motorways has been attributed by these authors to an increased mortality from roadkill, various degrees of habitat transformation, an increased effect of domestic carnivores, as well as increased habitat fragmentation and perturbation. On the other hand, positive associations between transport infrastructure and rabbit density in Spain have been related to the facilitation effects of verges for warren construction due to their increased soil depth (Barrientos & Bolonio 2009; Planillo & Malo 2017). In Malta, an average national soil depth of 48cm was reported by the MALSIS II project (Sultana 2017), with larger values found in cultivated areas comprising valley bottoms and sides that include rubble walls for soil erosion management. All sites observed in this study were located over the 'Garrigue' soil landscape type identified by the MALSIS project, characterized by a thin topsoil layer and karstic geology made of Upper Globigerina and Coralline limestone. Thus, the impervious roads measured in this study were unlikely to offer

motorway verges suitable for warren building, nor to allow the digging of tunnels under them as observed in Sweden by Ziege et al. (2020), which may explain the absence of positive effects from impervious roads due to facilitated warren construction along verges.

Moreover, the release of hunting pressure around motorway verges due to reduced frequentation of hunters along roads is unlikely in Malta. While it is illegal to hunt within 50m of main roads in Malta and a large part of the hunting is carried out at fixed locations, or 'trapping sites' (Borg 2002), there are known local enforcement issues of hunting regulations with little control on where it occurs (Ferns, Campbell & Verissimo 2021). Planillo and Malo (2017) found greater rabbit abundance associated with wider verges and increased vegetation cover, while it was lower at segments with medium to high traffic flows (>6000 vehicles/day). The Republic of Malta has one of the "densest [road] network in the world" with increasing traffic issues, spreading all over the country due to its small size that prevent large zones of very low network density as in mainland countries (Camilleri 2020, p.8). Hence, roads are present almost everywhere, and in those areas where hunting is heavily practiced, such as in the few remnants of open garrigue, their presence is unlikely to alter the hunting patterns.

Therefore, the observed negative effect of paved roads on rabbit density in the Maltese garrigue would more likely be linked to increased roadkill mortality, increased disturbance from domestic carnivores, and/or a more general habitat fragmentation effect hindering dispersal and colonisation between subpopulations. The fact that the road variable derived from the 200m buffer was retained, rather than the 500m buffer, suggests that the effect of motorways on rabbit density in the Maltese garrigue is likely due to proximity factors, which points to roadkill mortality and fragmentation.

Such negative effects seem also likely to be stronger than the potential beneficial effects of motorway verges for rabbits related to the vegetation, as noted elsewhere. Road verges in Malta have been found to possess similar vegetation communities to the garrigue, albeit with reduced richness and cover along surfaced roads (Ghose Roy 2023). This author also notes that road works tend to reduce verges' width along surfaced roads, which was positively related by Planillo and Malo (2017) to rabbit density, because motorway verges in Spain are typically accompanied by a fence placed at varying distances from the road itself. This is not often the case in Malta, where space is limited and roads are more often bordered by some kind of stone walls and narrow verges, when existent. Ghose Roy (2023) also states that the niche effect for the vegetation was more pronounced near permeable paths than surfaced roads.

The extent of permeable paths was the most important (negative) effect on the density of pellets identified in the current study. While only an approximate surface of paths was measured without any specific characteristics such as verge width, it is likely to still correlate with the availability of verges along paths. If such verges would offer a vegetation with a distinct effect on pellet density, through increased food or refuge availability, it should have been captured at least partially by the chosen variable. Therefore, the effect of vegetation along paths reported in the European rabbit's native range is unlikely to be significant in the Maltese garrigue. It does not necessarily imply that they are absent, but that instead, the negative effects associated with paths might be stronger, such as human disturbance and hunting pressure.

Maltese hunting is characterized by the ubiquitous use of trapping sites, as illustrated by Figure 5.2 below. These sites are interconnected by the type of permeable paths measured in this study and were indeed observed at most sites, which suggest that this variable may be associated with hunting pressure. Hunting and trapping sites are known to complicate access

to the public due to land tenure and management conflicts, deeply entrenched in the Maltese society (Pace 2022). This in turn causes excessive frequentation of the remaining open natural areas on weekends and during the tourist season (Pace 2022). Therefore, hiking paths and trails in Malta are associated either with intense hunting and trapping activities, or concentrated recreational activities in areas without hunting. The significant negative effect of permeable paths identified in this research is thus likely to indicate a strong effect of hunting and trapping, outdoor recreational activities, or both, on the spatial feeding patterns of the wild European rabbit in the Maltese garrigue. Rabbits appeared to have higher densities at sites with less trails in and around them up to 500m, which suggests that the effect of paths occurs at a larger scale than the effect of impervious roads. Ultimately, the negative effects of roads and paths suggest that their pressures on wild rabbit populations in the Maltese garrigue are more important than any positive effects they could induce, such as increased soil depth or reduced predation pressure.

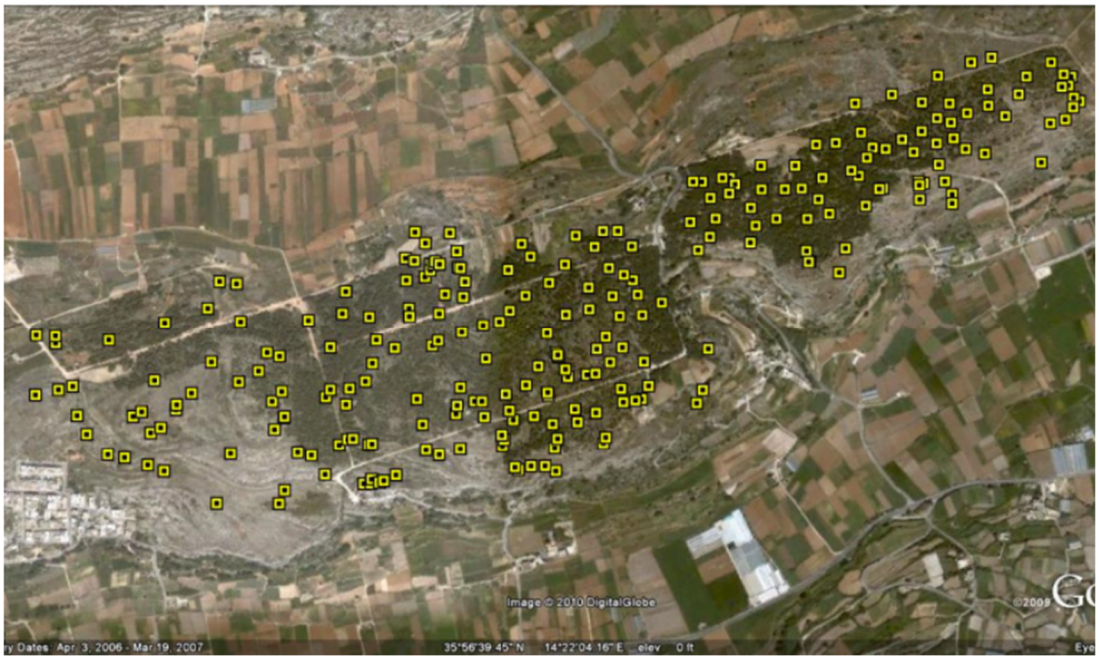


Figure 5.2. *An image from a 2010 report handed to MEPA/PA showing hunting and trapping structures in the Mizieb Woodland Nature Reserve. From Agius (2021, p. 145).*

5.4 Methodological considerations

The current dissertation used the compositional data analysis framework to study the effect of land cover variables derived by remote sensing. This framework stemmed from the work of Aitchison (1986) but was only generalized to applied sciences during the last two decades (Alenazi 2023). While it involves a steep initial learning curve due to the rather complex mathematics and ongoing academic debates, it proved useful to account for the relative nature of land covers measured inside a digital grid. Dividing space into regular shapes inside which the percentages of different land cover types is measured is nowadays common with the ubiquity of GIS in environmental sciences. Failing to account for the compositional nature of this type of data can lead to erroneous results and conclusions, including an increased risk of Type I/II errors. Not using log-ratio transformations to analyse compositional data can also result in reversed relationships, due to the misinterpretation of the proportional relationships and constraints inherent in the data when analysed in their raw form. Log-ratio transformations are essential for accurately representing this data in Euclidean space, thus avoiding misinterpretations that arise from the constant sum constraint. Here, it allowed to infer relationships between the land covers in addition to their relationships to the dependent variable. While it was harder to extract clear quantified effects from the models using log-ratios, they facilitated the interpretation of dynamics and trends. They allowed to distinctly observe and interpret the concurrent increase in herbaceous and rock covers against a simultaneous decrease in the three other land cover types. This specific pattern of change could then be directly associated with shifts in the garrigue landscape, transitioning from features more characteristic of steppes to those typical of maquis shrublands.

5.5 Limitations

This study was exploratory and therefore, all points mentioned above are rather speculative given the lack of existing local ecological knowledge on the European rabbit. Beside the limitations already mentioned in the discussion, a few points greatly constrain the results presented here.

The small sample size greatly reduced the statistical power of the whole analysis and the conclusions drawn from it (Campbell 2021). While some of the main garrigue areas in Malta were considered, they were not sufficiently sampled to allow a generalization to their whole extents or across the islands. Moreover, the procedure for site selection was not very robust and implied a strong subjective element, leading to a rather homogeneous selection. This was reflected in the low variability of some important variables, such as individual shrub height and area, restricted even further by the inclusion of a single habitat type. The observation of the selected variables across more diverse habitats, or along identified gradients, would allow to maximize variability in a more controlled manner and enhance the modelling procedure.

The distribution and the home range of the European rabbit in Malta is unknown, thus it was impossible to base the study design on known locations beside a rapid scouting during the pilot phase of the dissertation. While it somehow emulates a random selection, the effects of environmental and anthropogenic factors on rabbit density would be better observed based on gradients starting from previously identified locations, themselves part of the study design, such as warrens or latrines. Both have strong relationships to the territorial behaviour of the rabbit and represent natural population gradient centres (Sneddon 1991). When using free standing pellet counts such as in this study, it appears more appropriate to base gradients on some environmental variables instead, and sample pellets along these.

As mentioned above, shrub taxonomy was not considered, although contrasting effects were previously observed between different shrub families in Portugal (Beja et al. 2007). Due to the rabbit's highly adaptable behaviour, its relationship with local plant communities is complex and often includes divergent patterns depending on specific plant morphological and phytological characteristics. Similarly, various gradients in the observed land covers were not considered, such as soil hardness or depth. Variables without effects in this study might well mask contrasting effects of some of their components on rabbit density, such as a difference between species of shrubs.

The drone approach used for the vegetation survey implied a large initial labour investment, both for training and processing the data, as well as due to the numerous technical and logistic constraints, such as storage. Unfortunately, the processing logs were lost and could not be presented, but it raises the question of when to resort to higher technologies, and when simpler ones may be sufficient for a given task. A large sample size would have been difficult to handle with the given material configuration, even though it would have been beneficial to the study. On the other hand, more adequate equipment would have allowed faster survey and processing times. RTK modules allow to reduce or cancel the need for GCPs, which are labour intensive, while multispectral cameras facilitate the classification process, allowing to resort to faster classification algorithm such as RF. Nonetheless, the methodology developed here can be applied at lower costs with a commercial-grade RGB drone, even if it still relies on high-end GNSS equipment.

The methodology used for the classification of the land covers and the reclassification of refuge shrubs, low-lying shrubs and trees might have induced errors not quantified by the accuracy assessment. The robust regression applied on the AGMs was a crude way to account for unperfect data, due to the mismatch in resolution between the LiDAR (DTM) and the drone

(DSM) elevation datasets. This could encourage the use of a coarser resolution in drone surveys of the land covers in Malta to match the national LiDAR dataset's resolution (15cm GSD), in the absence of mobile LiDAR equipment for DTM acquisition. The current study used a 1cm GSD resolution to allow a precise characterisation of the shrubs' vertical structure (e.g. overhead canopy or dense base), however, it proved infeasible with the current survey parameters.

In fact, this ultrahigh-resolution approach was not fully leveraged while it involved high processing costs but few commensurate benefits to the work. In retrospect, two alternative drone survey approaches could have been considered to enhance efficiency and relevance to the study objectives. First, a larger-scale approach could involve flying at a higher altitude over more extensive areas, as mentioned above, which could have facilitated the observation of pellet distribution along greater variation gradients in the considered variables. This approach could offset the need for an initial scouting and the restriction to areas with proven rabbit presence only, because it would allow to consider zones with and without rabbits as part of a single dataset. Secondly, a smaller-scale, rabbit-centred approach could build more efficiently on an initial rabbit survey. Subsequent drone surveys could target specifically areas with a desired rabbit feature, such as warrens or latrines, and include only the surroundings in the drone survey. This method could offer a larger number of smaller sites and fully leverage the high-resolution imagery. Each of these alternative strategies underscores the importance of carefully considering the scale of drone surveys in ecological research. By aligning the survey's scale with the specific research questions and resource constraints, the potential benefits of drone technology can be maximized, ensuring that the data collected is both cost-effective and directly relevant to the study's objectives.

5.6 Implications for conservation and management

Thyme is yet another emblematic species of Malta, which provides glorious colours through the summer and a sought-after 'summer honey' on which the local beekeeping industry partly relies (Gambin, Lanfranco & Mifsud 2013). Its importance to the plant community ecology of the Maltese garrigue has been pointed out by Galea (2017). Here, it also appeared to be at the centre of the interactions between the considered land covers in the Maltese garrigue, while also being positively associated with wild European rabbit relative densities. The causality of the relationships was not elucidated, but any future project centred on the herbaceous layer or the low-lying shrub vegetation, including *Thymbra capitata*, should account for potential rabbit populations and their effects. It was unclear if the dynamic observed here was seasonal or continuous. Shrub encroachment is a phenomenon happening worldwide and heightened by climate change, although its consequences are complex and variables (Elridge et al. 2011). If the wild rabbit really favours shrub encroachment from thyme in the Maltese garrigue, it could interact with the effects of climate change to facilitate further shrub encroachment. However, the relevance of this potential change would still need to be evaluated for Malta.

This study supports the idea that anthropogenic features are a stronger determinant of the spatial distribution of the European rabbit in the Maltese garrigue than 'natural' features, albeit with the limitations mentioned in the next section of this chapter. This view is coherent with the long history of human occupation of the archipelago and its limited land mass, resulting in a prolonged and widespread artificialisation of the ecosystems (Schembri et al. 2009). In this context, the results reported here may be used to inform wild rabbit management in Malta, whether to conserve or to regulate it. Increasing wild populations in the garrigue would likely require the creation of areas without trails, hunting or trapping sites, and with minimal paved roads in and around them. This is a real challenge in the current Maltese context where

numerous conflicts for open space and its usages already exist (Agius 2021; Borg 2002; Falzon 2008). However, the interplay between *T. capitata*, herbaceous species and the European rabbit suspected in this study could have wider economic implications. If there is indeed a link between rabbit density and thyme in the Maltese garrigue, the honey industry could possibly find an interest in promoting areas of rabbit conservation and potential increased thyme cover in some of the open garrigues in Malta. Such areas could even be combined with permanent research plots for rabbit studies, as discussed below.

5.7 Future research directions

The exploratory design of this dissertation provided some insights for future research avenues on the European rabbit in Malta. Some important ecological metrics for wild *Oryctolagus cuniculus* are still missing locally. Notably, home-range size is an important parameter to correctly design any observational or experimental study of the wild rabbit, given its degree of plasticity. The use of free-standing pellets would be greatly enhanced from obtaining seasonal decay and accumulation rates in Malta, as well as by being related to absolute density indices.

Controlled experimental studies on the effects of the rabbit herbivory on the Maltese vegetation would provide causal insights, which are still missing entirely. This would be ideally complemented by diet studies, possibly using microhistological analyses of rabbit pellets related to community assemblages. Permanent plots with regular pellet clearance appear as a promising method to relate rabbit feeding behaviour to environmental parameters, in addition to allowing the observation of temporal effects as well.

Warren building behaviour of the local wild rabbit population have not been studied academically, while it has provided precious insights elsewhere. Unravelling the probably

interactive effects of geology, soil depth and hardness, rubble walls and soil accumulation, with rabbit warren density may allow to explain the lack of effects of the shrub cover observed in this study. Additionally, warren studies can provide information of the home range size and behaviour of the rabbit at the same time. The local rabbit hunting practices revolve around warren trapping with dogs and ferrets (Cassar 1994), and it is likely that the local hunting organisation (FKNK) concentrates a rich spatial and behavioural knowledge on rabbit warrens. A future collaboration with this organisation could help to connect conservation and recreational interests around the rabbit and open spaces in Malta.

As mentioned earlier, longitudinal studies are crucial to understand the dynamic of ecological relationships in general. In Malta, it would allow for the observation of the effect of the dry season and the related drastic changes in vegetation. One advantage of drone surveys is that they enable an increase in temporal resolution of the samples. Once a site is surveyed with the correct positioning equipment, and if permanent GCPs are created or readily available, it can be surveyed again without positioning equipment and accurately overlaid on the previous survey outputs. As such, drone surveys hold great potential for accurate repeated measurement of landscape features. This could be harnessed to study the effect of shrubs species across seasons on rabbits or follow the evolution of the herbaceous layer with and without rabbit grazing through time.

6. Conclusion

This dissertation quantified land covers and vegetation features in the Maltese garrigue using supervised SVM classification while anthropogenic features were derived from GIS and relative rabbit density was estimated through field-based quadrat survey. Compositional data analysis analysed land cover composition variations within fixed-dimension polygons. Ultimately, separate generalized linear models were used to identify the most influential factors in each set of explanatory variables.

Empirical results confirmed several educated conjectures about the main factors shaping *Oryctolagus cuniculus* habitat preferences in Malta's remaining open garrigue patches. Anthropogenic factors were more significant than environmental factors in explaining rabbits' spatial distribution at the dry season's onset. Paved roads and hiking trails negatively correlated with rabbit pellet density. The effect of paved roads was inferred to be related to increased mortality from roadkill and a fragmentation effect hindering colonisation and dispersal. The effect of hiking trails was inferred to be related to hunting pressure and human disturbance from recreational outdoor activities. Altogether, these outcomes suggest that the European rabbit in Malta thrives in more secluded areas, relatively to human disturbance, rather than selecting an optimal natural environment with regards to food or refuge availability.

Nonetheless, an effect of the land covers on rabbit density was detected, with the ratio of herbaceous to low-lying shrubs covers found to be negatively associated with rabbit density. The trade-off between food and refuge availability observed in the rabbit's native range was expected to result in a significant effect of the ratio between the herbaceous layer and higher shrubs, not the low-lying ones. The latter land cover mainly comprised *Thymbra capitata*, a dwarf shrub species endemic of the archipelago that grows and flowers through the summer dry season. At the onset of the dry season, rabbits were more abundant at sites that had a larger thyme cover than herbaceous cover in February. This was inferred to indicate that the seasonality of food availability might play a larger role than absolute nutritive value or palatability.

Overall, this study provided three main insights on the ecology of the wild European rabbit in the Maltese garrigue:

I. Human development and recreational activities significantly influence the spatial distribution of the wild European rabbit in Malta's remaining 'natural' habitats (Research question 3).
II. The vegetation dynamics of the Maltese garrigue seem to influence the spatial distribution of the wild European rabbit to a small extent, likely in relation to the surface of *Thymbra capitata* cover (Research question 2).
III. Human hunting pressure likely exceeds the impact of non-human predation on the spatial distribution of Malta's wild rabbits, because the role of shrubs for refuge did not appear to have an effect on the wild rabbit's distribution (Research question 1).

References

Abrantes, J, van der Loo, W, Le Pendu, J, & Esteves, PJ 2012, 'Rabbit haemorrhagic disease (RHD) and rabbit haemorrhagic disease virus (RHDV): a review', *Veterinary research*, vol. 43, no. 1, p. 12.

Agriculture Victoria 2021, *European rabbit*, viewed 14 July 2023, https://agriculture.vic.gov.au/biosecurity/pest-animals/priority-pest-animals/european-rabbit

Agius, K 2021, 'Ecotourism development in Malta: environmental issues and threats', in M Avellino & G Cassar (eds.), *The creative economy and sustainable tourism development – with a special focus on Malta,* Malta: Kite Group and University of Malta, Institute for Tourism, Travel and Culture.

Aitchison, J 1986, *The Statistical Analysis of Compositional Data*, Springer Netherlands, Dordrecht.

Akaike, H 1973, 'Information theory and an extension of the maximum likelihood principle', in BN Petrov and F Csaki, (eds), *International Symposium on Information Theory*, pp. 267-281.

Alenazi, A 2023, 'A review of compositional data analysis and recent advances', *Communications in Statistics - Theory and Methods*, vol. 52, no. 16, pp. 5535–5567.

Alrababah, MA, Alhamad, MA, Suwaileh, A & Al-Gharaibeh, M 2007, 'Biodiversity of semi-arid Mediterranean grasslands: Impact of grazing and afforestation', *Applied Vegetation Science*, vol. 10, pp. 257-264.

Altbäcker, V, Hudson, R, & Bilkó, Á 1995, 'Rabbit-mothers' diet influences pups' later food choice', *Ethology*, vol. 99, no. 1–2, pp. 107–116.

Anderson, K & Gaston, KJ 2013, 'Lightweight unmanned aerial vehicles will revolutionize spatial ecology', *Frontiers in Ecology and the Environment*, vol. 11, no. 3, pp. 138–146.

References

Arques, J, Belda, A, Peiro, V, Seva, E & Martinez-Perez, JE 2014, 'Main ecological gradients and landscape matrix affect wild rabbit *Oryctolagus cuniculus* (Linnaeus, 1758) (Mammalia: Leporidae) abundance in a coastal region (South-East Spain)', *Italian Journal of Zoology*, vol. 81, no. 3, pp. 440–450.

Aronne, G & Wilcock, CC 1994, 'Reproductive characteristics and breeding system of shrubs of the Mediterranean region', *Functional Ecology*, vol. 8, no. 1, p. 69.

Attard, G 2008, 'METIC agriculture: A resume of identified themes for collaboration', *Mediterranean Trading and Innovation Centre*, Valletta.

Bailey, G, Li, Y, McKinney, N, Yoder, D, Wright, W, & Herrero, H 2022, 'Comparison of ground point filtering algorithms for high-density point clouds collected by terrestrial LiDAR', *Remote Sensing*, vol. 14, no. 19, p. 4776.

Barbar, F & Lambertucci, SA 2018, 'The roles of leporid species that have been translocated: a review of their ecosystem effects as native and exotic species', *Mammal Review*, vol. 48, no. 4, pp. 245–260.

Basheer, S, Wang, X, Farooque, AA, Nawaz, RA, Liu, K, Adekanmbi, T, & Liu, S 2022, 'Comparison of land use land cover classifiers using different satellite imagery and machine learning techniques', *Remote Sensing*, vol. 14, no. 19, p. 4978.

Beja, P, Pais, M, & Palma, L 2007, 'Rabbit Oryctolagus Cuniculus Habitats in Mediterranean Scrubland: The Role of Scrub Structure and Composition', *Wildlife Biology*, vol. 13, no. 1, pp. 28–37.

Belgiu, M & Drăguț, L 2016, 'Random Forest in remote sensing: A review of applications and future directions', *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 114, pp. 24–31.

Benítez-López, A, Alkemade, R, & Verweij, PA 2010, 'The impacts of roads and other infrastructure on mammal and bird populations: A meta-analysis', *Biological Conservation*, vol. 143, no. 6, pp. 1307–1316.

Berhane, T, Lane, C, Wu, Q, Anenkhonov, O, Chepinoga, V, Autrey, B, & Liu, H 2017, 'Comparing pixel- and object-based approaches in effectively classifying wetland-dominated landscapes', *Remote Sensing*, vol. 10, no. 2, p. 46.

References

Birdlife 2019, Hunters now practically granted nine out of twelve months to hunt, viewed 23 February 2023, https://birdlifemalta.org/2019/05/hunters-now-practically-granted-nine-out-of-twelve-months-to-hunt/

Blank, L & Carmel, Y 2012, 'Woody vegetation patch types affect herbaceous species richness and composition in a Mediterranean ecosystem', *Community Ecology*, vol. 13, no. 1, pp. 72–81.

Bochet, E, Poesen, & Rubio, JL 2006, 'Runoff and soil loss under individual plants of a semi-arid Mediterranean shrubland: influence of plant morphology and rainfall intensity', *Earth Surf. Process. Landforms,* vol. 31, pp. 536–549.

Borg, R 2002, 'The evolution of hunting and trapping legislation in Malta', PhD thesis, Faculty of Law, University of Malta.

Boogaart, KG vd, Tolosana-Delgado, R & Bren, M 2006, 'Concepts for handling of zeros and missing values in compositional data', in E. Pirard (ed.), *Proceedings of the IAMG'2006 Annual Conference on "Quantitative Geology from multiple sources"*, Liege, Belgium.

Boogaart, KG vd & Tolosana-Delgado, R 2013, *Analyzing Compositional Data with R*, Springer Berlin Heidelberg, Berlin, Heidelberg, viewed on 29 August 2023, https://link.springer.com/10.1007/978-3-642-36809-7.

Brullo, S, Brullo, C, Cambria, S, & Giusso del Galdo, G 2020, *The Vegetation of the Maltese Islands*, Springer International Publishing, Cham.

Brullo, S & del Galdo, G 2006, 'Taxonomic remarks on the *Anthyllis hermanniae* L. (Fabaceae, Faboideae) species complex of the Mediterranean flora', *Novon*, vol. 16, pp. 304-314.

Cabrera-Rodriguez, F 2006, 'Microhabitat selection of the European rabbit on La Palma, Canary Islands, Spain', *Acta Theriologica*, vol. 51, no. 4, pp. 435–442.

Caldwell, MR & Klip, JMK 2022, 'Patterns of wildlife activity and predator-prey dynamics in a highly touristed area', *The Southwestern Naturalist*, vol. 66, no. 1, pp. 35-47.

Cameron, AC & Privedi, PK 2013, *Regression Analysis of Count Data*, 2nd edn, Cambridge University Press, New York.

References

Camilleri, K 2020, 'Research traffic congestion in Malta: An analysis of the attributes, attitudes and solutions', *MCAST Journal of Applied Research & Practice*, vol. 4, no. 1, pp. 4-34.

Carvalho, JC & Gomes, P 2004, 'Influence of herbaceous cover, shelter and land cover structure on wild rabbit abundance in NW Portugal', *Acta Theriologica*, vol. 49, no. 1, pp. 63–74.

Cassar, C 1994, *Fenkata: an emblem of Maltese peasant resistance?,* Ministry for Youth and the Arts, Malta.

Cassar, LF, Conrad, E, & Schembri, PJ 2008, 'The Maltese Archipelago', in I Vogiatzakis, G Pungetti, & AM Mannion (eds), *Mediterranean Island Landscapes, Landscape Series*, pp. 297–322, Springer Netherlands, Dordrecht.

Cauchi, D 2014, 'The influence of habitat fragmentation: A study on the indigenous Least Weasel (*Mustela nivalis*) population in Malta', B.Sc. dissertation, University of Malta.

Cerdà, A, Lucas-Borja, ME, Franch-Pardo, I, Úbeda, X, Novara, A, López-Vicente, M, Popović, Z, & Pulido, M 2021, 'The role of plant species on runoff and soil erosion in a Mediterranean shrubland', *Science of The Total Environment*, vol. 799, p. 149218.

Chatzimpaloglou, P, Schembri, PJ, French, C, Ruffell, A, & Stoddart, S 2020, 'Chapter 1: the geology, soils and present-day environment of Gozo and Malta', in C French, CO Hunt, R Grima, R McLaughlin, S Stoddart & C Malone (eds*), Temple landscapes: fragility, change and resilience of Holocene environments in the Maltese Islands*, pp. 19-33, McDonald Institute for Archaeological Research, Cambridge

Cingolani, AM, Noy-Meir, I, & Díaz, S 2005, 'Grazing effects on rangeland diversity: A synthesis of contemporary models', *Ecological Applications*, vol. 15, no. 2, pp. 757–773.

Coffin, AW 2007, 'From roadkill to road ecology: A review of the ecological effects of roads', *Journal of Transport Geography*, vol. 15, no. 5, pp. 396–406.

Cook, B 2008, 'Managing the European rabbit: converging interests between Australian research for rabbit control and European research for their conservation', in PC Alves, N Ferrand & K Hackländer (eds), *Lagomorph biology: evolution, ecology and conservation*, pp 317–326, Springer, Berlin.

Courchamp, F, Chapuis, J-L, & Pascal, M 2003, 'Mammal invaders on islands: impact, control and control impact', *Biological Reviews*, vol. 78, no. 3, pp. 347–383.

Cubas, J, Martín-Esquivel, JL, Nogales, M, Irl, SDH, Hernández-Hernández, R, López-Darias, M, Marrero-Gómez, M, del Arco, MJ, & González-Mancebo, JM 2017, 'Contrasting effects of invasive rabbits on endemic plants driving vegetation change in a subtropical alpine insular environment', *Biological invasions*, vol. 20, no. 3, pp. 793–807.

De Giglio, M, Greggio, N, Goffo, F, Merloni, N, Dubbini, M, & Barbarella, M 2019, 'Comparison of pixel- and object-based classification methods of Unmanned Aerial Vehicle data applied to coastal dune vegetation communities: Casal Borsetti case study', *Remote Sensing*, vol. 11, no. 12, p. 1416.

Dellafiore, CM, Gallego Fernández, JB, & Vallés, SM 2008, 'Habitat use for warren building by European rabbits (*Oryctolagus cuniculus*) in relation to landscape structure in a sand dune system', *Acta Oecologica*, vol. 33, no. 3, pp. 372–379.

Delibes-Mateos, M, Redpath, SM, Angulo, E, Ferreras, P, & Villafuerte, R 2007, 'Rabbits as a keystone species in southern Europe', *Biological Conservation*, vol. 137, no. 1, pp. 149–156.

Delibes-Mateos, M, Smith, AT, Slobodchikoff, CN, & Swenson, JE 2011, 'The paradox of keystone species persecuted as pests: A call for the conservation of abundant small mammals in their native range', *Biological Conservation*, vol. 144, no. 5, pp. 1335–1346.

Delibes-Mateos, M, Rödel, HG, Rouco, C, Alves, PC, Carneiro, M & Villafuerte, R 2021, 'European Rabbit *Oryctolagus cuniculus* (Linnaeus, 1758)', in K Hackländer, FE Zachos (eds), *Handbook of the Mammals of Europe*, Springer, Cham.

Díaz Barradas, MC, Zunzunegui, M, Tirado, R, Ain-Lhout, F, & García Novo, F 1999, 'Plant functional types and ecosystem function in Mediterranean shrubland', Journal of Vegetation Science, vol. 10, no. 5, pp. 709–716.

Dunn, PK & Smyth, GK 2018, *Generalized Linear Models with examples in R*, Springer New York, New York, NY.

EASA 2009, Policy statement: airworthiness certification of unmanned aircraft systems, viewed 25 November 2022,

https://www.easa.europa.eu/sites/default/files/dfu/E.Y013-01_%20UAS_%20Policy.pdf

Eldridge, DJ, Bowker, MA, Maestre, FT, Alonso, P, Mau, RL, Papadopoulos, J, & Escudero, A 2010, 'Interactive Effects of Three Ecosystem Engineers on Infiltration in a Semi-Arid Mediterranean Grassland', *Ecosystems*, vol. 13, no. 4, pp. 499–510.

Elzinga, CL, Salzer, DW, Willoughby, JW & Gibbs, JP 2001, *Monitoring plant and animal populations*, Blackwell Science, Malden.

ERA 2018, *State of the Environment report.* Available from: https://www.parlament.mt/media/98024/dok-42-l11.pdf

ESRI n.d. (a), Generalization of classified raster imagery, viewed 27 March 2023, https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-analyst/generalization-of-classified-raster-imagery.htm

ESRI n.d. (b), How Average Nearest Neighbor works, viewed 13 February 2023, https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-statistics/h-how-average-nearest-neighbor-distance-spatial-st.htm

Ezzat Ahmed, A, Alkahtani, M, & Abdel-Wareth, A 2020, 'Thyme leaves as an eco-friendly feed additive improves both the productive and reproductive performance of rabbits under hot climatic conditions', *Veterinární medicína*, vol. 65, no. 12, pp. 553–563.

Falzon, MA 2008, 'Flights of passion: Hunting, ecology and politics in Malta and the Mediterranean', *Anthropology Today*, vol. 24, no. 1, pp. 15–20.

Faraway, J 2002, *Practical regression and ANOVA using R*, viewed 10 October 2023, https://cran.r-project.org/doc/contrib/Faraway-PRA.pdf

Farrell, M, Hunt, CO & McClung, L 2020, 'Chapter 3: the Holocene vegetation history of the Maltese Islands', in C French, CO Hunt, R Grima, R McLaughlin, S Stoddart & C Malone (eds*), Temple landscapes: fragility, change and resilience of Holocene environments in the Maltese Islands*, pp. 73-113, McDonald Institute for Archaeological Research, Cambridge.

Fernandes, P 2012, 'Study of a wild rabbit (*Oryctolagus cuniculus*) population: Habitat factors related to its spatial distribution', M.Sc. thesis, University of Coimbra.

Fernández, N 2005, 'Spatial Patterns in European Rabbit Abundance After a Population Collapse', *Landscape Ecology*, vol. 20, no. 8, pp. 897–910.

Fernandez-de-Simon, J, Díaz-Ruiz, F, Cirilli, F, Tortosa, FS, Villafuerte, R, Delibes-Mateos, M, & Ferreras, P 2011, 'Towards a standardized index of European rabbit abundance in Iberian Mediterranean habitats', *Eur J Wildl Res*, vol. 57, no. 5, pp. 1091–1100.

Ferns, B, Campbell, B & Verissimo, D 2021, 'Emerging contradictions in the enforcement of bird hunting regulations in Malta', *Conservation Science and Practice*, vol. 4, no. e12655, pp. 1-13.

Ferrand, N & Branco, M 2007, 'The evolutionary history of the European rabbit (*Oryctolagus cuniculus*): major patterns of population differentiation and geographic expansion inferred from protein polymorphism', in S. Weiss & N. Ferrand (eds), *Phylogeography of Southern European Refugia*, pp. 207-235, Springer Netherlands, Dordrecht.

Flux, JE. & Fullagar, PJ 1992, 'World distribution of the rabbit Oryctolagus cuniculus on islands', Mammal review, vol. 22, no. 3–4, pp. 151–205.

Forman, RTT & Alexander, LE 1998, 'Roads and their major ecological effects', *Annual Review of Ecology and Systematics*, vol. 29, no. 1, pp. 207–231.

Fromin, N, Shihan, A, Santonja, M, Baldy, V, & Hättenschwiler, S 2020, 'Soil microbial activity in a Mediterranean garrigue responds more to changing shrub community than to reduced rainfall', *Plant and Soil*, vol. 449, no. 1–2, pp. 405–421.

Furlani, S, Antonioli, F, Gambin, T, Biolchi, S, Formosa, S, Lo Presti, V, Mantovani, M, Anzidei, M, Calcagnile, L & Quarta, G 2018, 'Submerged speleothem in Malta indicates tectonic stability throughout the Holocene', *The Holocene*, vol. 28, no. 10, pp. 1588-1597.

Galdies, C, Betts, JC, Vassallo, A & Micallef, A 2014, 'High resolution agriculture land cover using aerial digital photography and GIS: A case study for small island states', in S Formosa (ed), *Future Preparedness: Thematic and Spatial Issues for the Environment and Sustainability,* pp.127-151, University of Malta, Malta.

Galdies, C 2022, *The State of the Climate 2022 - a multidecadal report and assessment of Malta's climate,* National Statistics Office, Malta.

Galea, L 2017, 'Multiscale analysis of the species diversity of the Maltese shrubland', M.Sc dissertation, Department of Biology, University of Malta.

Gálvez Bravo, L, Belliure, J, & Rebollo, S 2009, 'European rabbits as ecosystem engineers: warrens increase lizard density and diversity', *Biodiversity and Conservation*, vol. 18, no. 4, pp. 869–885.

Gálvez, L, López-Pintor, A, De Miguel, JM, Alonso, G, Rueda, M, Rebollo, S, & Gómez-Sal, A 2008, 'Ecosystem engineering effects of European rabbits in a Mediterranean habitat', in PC Alves, N Ferrand, & K Hackländer (eds), *Lagomorph Biology*, pp.125–139, Springer Berlin Heidelberg, Berlin.

Gambin, C, Lanfranco, E, & Mifsud, D 2013, 'Pollen characterisation of Maltese honey', *Xjenza Online*, vol. 1, no. 2, pp. 33-46.

Gao, J & Carmel, Y 2020a, 'Can the intermediate disturbance hypothesis explain grazing–diversity relations at a global scale?', *Oikos*, vol. 129, no. 4, pp. 493–502.

Gao, J & Carmel, Y 2020b, 'A global meta-analysis of grazing effects on plant richness', *Agriculture, Ecosystems & Environment*, vol. 302, p. 107072.

Gatt, AD 2021, 'Malta's drone innovation ecosystem', Times of Malta, viewed 19 December 2023, https://timesofmalta.com/articles/view/maltas-drone-innovation-ecosystem.917806

Gauci, A, Abela, J, Austad, M, Cassar, LF & Zarb Adami, K 2018, 'A Machine Learning approach for automatic land cover mapping from DSLR images over the Maltese Islands', *Environmental Modelling & Software*, vol 99, pp.1-10.

Gerlach, G & Musolf, K 2000, 'Fragmentation of Landscape as a Cause for Genetic Subdivision in Bank Voles', *Conservation Biology*, vol. 14, no. 4, pp. 1066–1074.

Ghose Roy, R 2023, 'Diversity of roadside flora in the Maltese Islands', M.Sc. thesis, Department of Biology, University of Malta.

Gislason, PO, Benediktsson, JA, & Sveinsson, JR 2006, 'Random Forests for land cover classification', *Pattern Recognition Letters*, vol. 27, no. 4, pp. 294–300.

Greenacre, M 2021, 'Compositional data analysis', *Annu. Rev. Stat. Appl.*, vol. 8, pp. 271–299

Greenacre, M 2019, *Compositional Data Analysis in Practice*, CRC Press, Boca Raton.

Hardy, C, Vigne, J-D, Casane, D, Dennebouy, N, Mounolou, J-C, & Monnerot, M 1994, 'Origin of European rabbit (Oryctolagus cuniculus) in a Mediterranean island: Zooarchaeology and ancient DNA examination', *Journal of Evolutionary Biology*, vol. 7, no. 2, pp. 217–226.

Hartig, F 2016, *DHARMa: residual diagnostics for hierarchical (multi-level / mixed) regression models*, R package version 0.1.0.

Hellicar, MA & Kirschel, ANG 2022, 'Influence of grazing and fire on breeding birds and perennial plants in Cyprus scrub and forest systems', *Journal for Nature Conservation*, vol. 68, p. 126207.

Hilbe, JM 2011, *Negative binomial regression*, 2nd eds, Cambridge University Press, United Kingdoms.

Hirakawa, H 2001, 'Coprophagy in leporids and other mammalian herbivores', *Mammal Review*, vol. 31, no. 1, pp. 61–80.

Hopkins, HL & Kennedy, ML 2004, 'An assessment of indices of relative and absolute abundance for monitoring populations of small mammals', *Wildlife Society Bulletin*, vol. 32, no. 4, p. 1289.

Hossain, MD & Chen, D 2019, 'Segmentation for Object-Based Image Analysis (OBIA): A review of algorithms and challenges from remote sensing perspective', *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 150, pp. 115–134.

Hron, K, Coenders, G, Filzmoser, P, Palarea-Albaladejo, J, Faměra, M, & Matys Grygar, T 2021, 'Analysing Pairwise Logratios Revisited', *Mathematical Geosciences*, vol. 53, no. 7, pp. 1643–1666.

Hunt, CO & Schembri, PJ 1999, 'Quaternary environments and biogeography of the Maltese Islands', in A Mifsud & C Savona Ventura (eds), *Facets of Maltese prehistory*, pp. 41-75, The Prehistoric Society of Malta, Malta.

Iborra, O & Lumaret, JP 1997, 'Validity limits of the pellet group counts in wild rabbit (*Oryctolagus cuniculus*)', *Mammalia,* vol. 61, no. 2, pp. 205–218.

References

ICAO 2011, Unmanned aircraft systems, viewed 25 November 2022, https://www.icao.int/meetings/uas/documents/circular%20328_en.pdf

Kalacska, M, Lucanus, O, Arroyo-Mora, J, Laliberté, É, Elmer, K, Leblanc, G, & Groves, A 2020, 'Accuracy of 3D landscape reconstruction without ground control points using different UAS platforms', *Drones*, vol. 4, no. 2, p. 13.

Keeley, JE 1986, 'Resilience of mediterranean shrub communities to fires', in B Dell, AJM Hopkins, & BB Lamont (eds), *Resilience in mediterranean-type ecosystems*, *Tasks for vegetation science*, pp.95–112, Dordrecht, Springer Netherlands.

Kontsiotis, VJ, Bakaloudis, DE, & Liordos, V 2018, 'Impact of European wild rabbits foraging in different habitat and vegetation types in an insular environment', *Mammalia*, vol. 82, no. 2, pp. 193–196.

Lanfranco, GG 1969, *Maltese mammals*, Progress Press, Malta.

Lang, DM 1960, *Soils of Malta and Gozo*, Colonial Research Series vol. 29, HMSO, London.

Lees, AC & Bell, DJ 2008, 'A conservation paradox for the 21st century: The European wild rabbit *Oryctolagus cuniculus*, an invasive alien and an endangered native species', *Mammal Review*, vol. 38, no. 4, pp. 304–320.

Legendre, P & Legendre, L 2012, *Numerical ecology,* Elsevier, Amsterdam.

Liu, X, Lian, X, Yang, W, Wang, F, Han, Y, & Zhang, Y 2022, 'Accuracy assessment of a UAV direct georeferencing method and impact of the configuration of ground control points', *Drones*, vol. 6, no. 2, p. 30.

Lombardo, E, Bancheva, S, Domina, G, & Venturella, G 2020, 'Distribution, ecological role and symbioses of selected shrubby species in the Mediterranean Basin: a review', *Plant Biosystems - An International Journal Dealing with all Aspects of Plant Biology*, vol. 154, no. 4, pp. 438–454.

Lopez-Martinez, N 2008, 'The Lagomorph fossil record and the origin of the European rabbit', in PC Alves, N Ferrand & K Hackländer (eds), *Lagomorph Biology: evolution, ecology and conservation*, pp 27–46, Springer, Berlin.

Lowe, DG 2004, 'Distinctive image features from scale-invariant keypoints', *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110.

Lu, D & Weng, Q 2007, 'A survey of image classification methods and techniques for improving classification performance', *International Journal of Remote Sensing*, vol. 28, no. 5, pp. 823–870.

Marín-García, PJ & Llobat, L 2021, 'What are the keys to the adaptive success of European wild rabbit (*Oryctolagus cuniculus*) in the Iberian Peninsula?', *Animals*, vol. 11, no. 8, p. 2453.

Martin-Fernandez, JA, Barcelo-Vidal, C, & Pawlowsky-Glahn, V 2003, 'Dealing with zeros and missing values in compositional data sets using nonparametric imputation', *Mathematical Geology*, vol. 35, no. 3, pp. 253-278

Martins, H, Barbosa, H, Hodgson, M, Borralho, R, & Rego, F 2003, 'Effect of vegetation type and environmental factors on European wild rabbit *Oryctolagus cuniculus* counts in a southern Portuguese montado', *Acta Theriologica*, vol. 48, no. 3, pp. 385–398.

Martins, H, Milne, JA, & Rego, F 2002, 'Seasonal and spatial variation in the diet of the wild rabbit (Oryctolagus cuniculus L.) in Portugal', *Journal of Zoology*, vol 258, pp. 395-404.

Masseti, M & De Marinis, AM 2008, 'Prehistoric and historic artificial dispersal of lagomorphs on the Mediterranean islands', in PC Alves, N Ferrand & K Hackländer (eds), *Lagomorph biology: evolution, ecology and conservation*, pp 13–25, Springer, Berlin.

Maulik, U & Chakraborty, D 2017, 'Remote Sensing image classification: A survey of Support-Vector-Machine-based advanced techniques', IEEE Geoscience and Remote Sensing Magazine, vol. 5, no. 1, pp. 33–52.

Maxwell, SL, Fuller, RA, Brooks, TM & Watson, JEM 2016, 'Biodiversity: The ravages of guns, nets and bulldozers', *Nature*, vol. 536, pp. 143-145.

Milchunas, DG, Sala, OE, & Lauenroth, WK 1988, 'A Generalized Model of the Effects of Grazing by Large Herbivores on Grassland Community Structure', The American Naturalist, vol. 132, no. 1, pp. 87–106.

Monzon, A, Fernandes, P, & Rodrigues, N 2004, 'Vegetation structure descriptors regulating the presence of wild rabbit in the National Park of Peneda-Geres, Portugal', *European Journal of Wildlife Research*, vol. 50, no. 1, pp. 1–6.

Mountrakis, G, Im, J, & Ogole, C 2011, 'Support vector machines in remote sensing: A review', *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 66, no. 3, pp. 247–259.

Muller, I, Hron, K, Fiserova, E, Smahaj, J, Cakirpaloglu, P, & Vancakova, J 2018, 'Interpretation of compositional regression with application to time budget analysis', *Austrian Journal of Statistics*, vol. 47, no. 2, pp. 3–19.

Mutze, G, Cooke, B, Lethbridge, M, & Jennings, S 2014, 'A rapid survey method for estimating population density of European rabbits living in native vegetation', *The Rangeland journal*, vol. 36, no. 3, p. 239.

Muscat, M 2021, 'Wild rabbits enduring pain and suffering before dying', TVM, viewed 13 November 2021, https://www.tvm.com.mt/en/news/wild-rabbits-enduring-pain-and-suffering-before-dying/

Mykytowycz, R & Gambale, S 1965, 'A study of the inter-warren activities and dispersal of wild rabbits, *Oryctolagus cuniculus* (L.), living in a 45-AC paddock', *CSIRO Wildlife Research*, vol. 10, no. 1, p. 111-123.

Nelder, JA & Wedderburn, RWM 2023, 'Generalized Linear Models', *Journal of the Royal Statistical Society,* vol. 135, no. 3, pp. 370-384.

Nowak, MM, Dziób, K, & Bogawski, P 2019, 'Unmanned Aerial Vehicles (UAVs) in environmental biology: a review', *European Journal of Ecology*, vol. 4, no. 2, pp. 56–74.

Pace, D 2022, 'Tour Guides and Access to Trails: Problems in the Baħrija Area of Malta', *International Journal of Tour Guiding Research*, vol. 3, no. 1, pp. 1-18.

Palomares, F 2001, 'Comparison of 3 Methods to Estimate Rabbit Abundance in a Mediterranean Environment', *Wildlife Society Bulletin*, Vol. 29, No. 2, pp. 578-585.

Pawlowsky-Glahn, V, Egozcue, JJ & Tolosana-Delgado, R 2015, *Modeling and Analysis of Compositional Data*, John Wiley & Sons, Chichester.

Pescador, D, de la Cruz, M, Chacon-Labella, J, Pavon-Garcia, J & Escudero, A 2020, 'Tales from the underground: soil heterogeneity and not only above-ground plant interactions explain fine-scale species patterns in a Mediterranean dwarf shrubland', *Journal of Vegetation Science*, vol. 31, pp. 497-508.

Pijl, A, Bailly, J-S, Feurer, D, El Maaoui, MA, Boussema, MR, & Tarolli, P 2020, 'TERRA: Terrain Extraction from elevation Rasters through Repetitive Anisotropic filtering', *International Journal of Applied Earth Observation and Geoinformation*, vol. 84, p. 101977.

Planillo, A & Malo, JE 2017, 'Infrastructure features outperform environmental variables explaining rabbit abundance around motorways', *Ecology and Evolution*, vol. 8, no. 2, pp. 942–952.

Pol, RG, Sagario, MC, & Marone, L 2014, 'Grazing impact on desert plants and soil seed banks: Implications for seed-eating animals', *Acta Oecologica*, vol. 55, pp. 58–65.

Putman, RJ 1984, 'Facts from faeces', *Mammal review*, vol. 14, no. 2, pp. 79–97.

Rahmanian, S, Ejtehadi, H, Farzam, M, Hejda, M, Memariani, F, & Pyšek, P 2021, 'Does the intensive grazing and aridity change the relations between the dominant shrub Artemisia kopetdaghensis and plants under its canopies?', *Ecology and Evolution*, vol. 11, no. 20, pp. 14115–14124.

R Core Team 2023, *R: A language and environment for statistical computing*, R Foundation for Statistical Computing, Vienna, Austria, https://www.R-project.org/.

Richards, JA 2022, *Remote Sensing Digital Image Analysis*, Springer International Publishing, Cham.

Rodrigues, M, Bos, AR, Schembri, PJ, de Lima, RF, Lymberakis, P, Parpal, L, Cento, M, Ruette, S, Ozkurt, SO, Santos-Reis, M, Merilä, J, & Fernandes, C 2017, 'Origin and introduction history of the least weasel (*Mustela nivalis*) on Mediterranean and Atlantic islands inferred from genetic data', *Biological Invasions*, vol. 19, no. 1, pp. 399–421.

Salach, A, Bakuła, K, Pilarska, M, Ostrowski, W, Górski, K, & Kurczyński, Z 2018, 'Accuracy Assessment of Point Clouds from LiDAR and Dense Image Matching Acquired Using the UAV Platform for DTM Creation', *ISPRS International Journal of Geo-Information*, vol. 7, no. 9, p. 342.

Sanz-Ablanedo, E, Chandler, J, Rodríguez-Pérez, J, & Ordóñez, C 2018, 'Accuracy of Unmanned Aerial Vehicle (UAV) and SfM photogrammetry survey as a function of the number and location of ground control points used', *Remote Sensing*, vol. 10, no. 10, p. 1606.

Schembri, PJ, Hunt, C, Pedley, M, Malone, C, & Stoddart, S 2009,' The environment of the Maltese islands', in C Malone, S Stoddart, A Bonanno, D Trump, T Gouder & A Pace (eds), *Mortuary customs in prehistoric Malta: excavations at the Brochtorff Circle at Xagħra (1987-94)*, pp. 17-39, McDonald Institute for Archaeological Research Cambridge.

Schembri, PJ 1997, 'The Maltese Islands: climate, vegetation and landscape', *GeoJournal*, vol. 41, no. 2, pp. 115-125.

Schembri, PJ & Lanfranco, E 1996, 'Introduced species in the Maltese Islands', in AE Baldacchino, & A Pizzuto (eds), *Introduction of alien species of flora and fauna*, pp.29-54, Malta: Environment Protection Department, Floriana.

Schembri PJ 1993, 'Physical geography and ecology of the Maltese Islands: a brief overview', in S Busuttil, F Lerin, & L Mizzi (eds), *Malta: food, agriculture, fisheries and the environment*, pp. 27-39, CIHEAM, Montpellier.

Schembri, P J, & Lanfranco, E 1993, 'Development and the natural environment in the Maltese Islands', in D G Lockhart, D Drakakis-Smith, & J Schembri (eds), *The development process in small island states*, pp. 247-266, Routledge, London.

Schweizer, D, Jones, HP, & Holmes, ND 2016, 'Literature review and meta-analysis of vegetation responses to goat and European rabbit eradications on islands', *Pacific Science*, vol. 70, no. 1, pp. 55–71.

Serronha, AM 2014, 'Modeling the factors limiting the distribution and abundance of the European rabbit (*Oryctolagus cuniculus)* in SE Portugal', M.Sc. dissertation, Instituto Superior de Agronomia, Universidade de Lisboa.

Sibaruddin, HI, Shafri, HZM, Pradhan, B, & Haron, NA 2018, 'Comparison of pixel-based and object-based image classification techniques in extracting information from UAV imagery data', *IOP Conference Series: Earth and Environmental Science*, vol. 169, p. 012098.

Sneddon, IA 1991, 'Latrine Use by the European Rabbit (*Oryctolagus cuniculus*)', *Journal of mammalogy*, vol. 72, no. 4, pp. 769–775.

Rey, C & Saez, F 2002, 'Field culture, *in vitro* culture and selection of *Thymus',* in E Stahl-Biskup & F Saez (eds), *Thyme: the genus Thymus*, Taylor & Francis, London, pp. 177-196.

Silva, C, Simões, MP, Mira, A, & Santos, SM 2019, 'Factors influencing predator roadkills: The availability of prey in road verges', *Journal of Environmental Management*, vol. 247, pp. 644–650.

Stoklosa, J, Blakey, RV, & Hui, FKC 2022, 'An Overview of Modern Applications of Negative Binomial Modelling in Ecology and Biodiversity', *Diversity*, vol. 14, no. 320, pp. 1-25.

Štroner, M, Urban, R, Lidmila, M, Kolář, V, & Křemen, T 2021, 'Vegetation Filtering of a Steep Rugged Terrain: The Performance of Standard Algorithms and a Newly Proposed Workflow on an Example of a Railway Ledge', *Remote Sensing*, vol. 13, no. 15, p. 3050.

Sultana, D 2017, 'Soil Quality Change in the Maltese Islands: A 10-Year Assessment (2003 to 2013)', in S Formosa (ed), *Emergent realities for social wellbeing: environmental, spatial and social pathways*, pp.267-292, University of Malta, Malta.

Sultana, D 2015, 'Numerical modelling of soil erosion susceptibility in the Maltese islands using geographic information systems and the revised universal soil loss equation (RUSLE)', *Xjenza Online*, vol. 3, pp. 41–50.

Tranchant, J 2022, 'An appraisal of the *Oryctolagus cuniculus* population on the island of Comino and its relationship to the vegetation', B.Sc. dissertation, Institute of Earth Systems, University of Malta.

Ulvi, A 2021, 'The effect of the distribution and numbers of ground control points on the precision of producing orthophoto maps with an unmanned aerial vehicle', *Journal of Asian Architecture and Building Engineering*, vol. 20, no. 6, pp. 806–817.

Vella, SJ 2005, 'Soil survey and soil mapping in the Maltese islands: the 2003 position', in RJA Jones, B Houšková, P Bullock & L Montanarella (eds), *Soil Resources of Europe, second edition*, pp. 235-244, Office for Official Publications of the European Communities, Luxembourg.

Veterinary Regulation Directorate 2021, viewed 27 July 2023, https://agrikoltura.gov.mt/en/vrd/Pages/home.aspx

Villafuerte, R & Delibes-Mateos, M 2019, '*Oryctolagus cuniculus* (errata version published in 2020)', *The IUCN Red List of Threatened Species* 2019: e.T41291A170619657, viewed 27 December 2022, https://www.iucnredlist.org/species/41291/170619657

Vogiatzakis, I, & Griffiths, GH 2008, 'Island Biogeography and Landscape Ecology', in I Vogiatzakis, G Pungetti, & AM Mannion (eds), *Mediterranean Island Landscapes, Landscape Series*, pp. 61–81, Springer Netherlands.

Whiteside, TG, Boggs, GS, & Maier, SW 2011, 'Comparing object-based and pixel-based classifications for mapping savannas', *International Journal of Applied Earth Observation and Geoinformation*, vol. 13, no. 6, pp. 884–893.

Willott, SJ, Miller, AJ, Incoll, LD, & Compton, SG 2000, 'The contribution of rabbits (*Oryctolagus cuniculus* L.) to soil fertility in semi-arid Spain', *Biology and Fertility of Soils*, vol. 31, no. 5, pp. 379–384.

Yu, JJ, Kim, DW, Lee, EJ, & Son, SW 2020, 'Determining the optimal number of ground control points for varying study sites through accuracy evaluation of unmanned aerial system-based 3D point clouds and digital surface models', *Drones*, vol. 4, no. 3, p. 49.

Ziege, M, Hermann, BT, Kriesten, S, Merker, S, Ullmann, W, Streit, B, Wenninger, S, & Plath, M 2020, 'Ranging behaviour of European rabbits (*Oryctolagus cuniculus*) in urban and suburban landscapes', *Mammal Research*, vol. 65, no. 3, pp. 607–614.

Zeileis, A, Kleiber, C, & Jackman, S 2008, 'Regression Models for Count Data in *R*', *Journal of Statistical Software*, vol. 27, no. 8, pp. 1-25.

Appendices

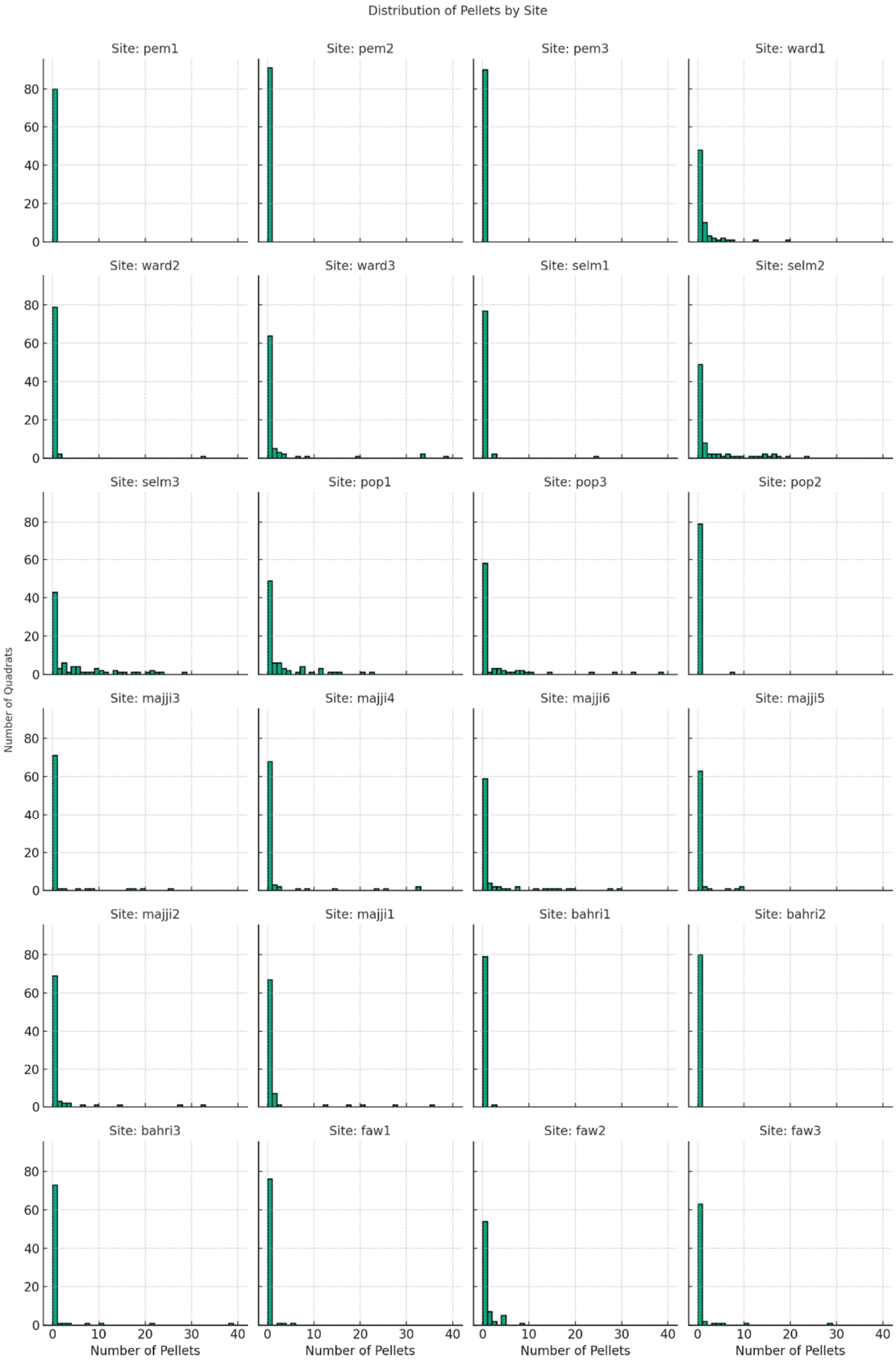
Appendix 1. Complete data tables

A1.1. Dung pellets survey results

All sites had 80 quadrats surveyed, equalling 8m² surveyed per 1ha site.

Area of Interest	Site	Quadrats on latrines	Total pellets outside latrines
Bahrija	bahri1	0	2
Bahrija	bahri2	0	0
Bahrija	bahri3	4	82
Il-Fawwara	faw1	3	10
Il-Fawwara	faw2	1	39
Il-Fawwara	faw3	3	52
Il-Majjistral	majji1	2	120
Il-Majjistral	majji2	3	101
Il-Majjistral	majji3	0	100
Il-Majjistral	majji4	2	147
Il-Majjistral	majji5	2	36
Il-Majjistral	majji6	4	199
Pembroke	pem1	0	0
Pembroke	pem2	0	0
Pembroke	pem3	0	0
Popeye Village	pop1	3	195
Popeye Village	pop2	0	7
Popeye Village	pop3	2	219
Selmun	selm1	1	28
Selmun	selm2	3	237
Selmun	selm3	3	358
Wardija	ward1	4	80
Wardija	ward2	1	34
Wardija	ward3	3	154

A1.2. Dung pellets statistical distribution by site



A1.3. Shapiro-Wilk normality test results for pellet counts by site

Site	SW Statistic	p-value
bahri1	0.09	0
bahri3	0.22	0
faw1	0.18	0
faw2	0.46	0
faw3	0.21	0
majji1	0.29	0
majji2	0.27	0
majji3	0.32	0
majji4	0.32	0
majji5	0.29	0
majji6	0.48	0
pop1	0.59	0
pop2	0.09	0
pop3	0.45	0
selm1	0.11	0
selm2	0.61	0
selm3	0.70	0
ward1	0.44	0
ward2	0.10	0
ward3	0.30	0

A1.2. Shrub metrics by site

Site name	Number of shrubs	Average shrub area (m^2)	Average shrub maximum height (m)	Average shrub height (m)	Nearest Neighbour index	Average distance to nearest neighbour (m)
Bahrija 1	1303	0.69	0.38	0.26	0.98	1.05
Bahrija 2	571	0.64	0.37	0.24	0.78	0.97
Bahrija 3	295	0.64	0.35	0.24	0.80	1.41
Il-Fawwara 1	2139	0.71	0.40	0.26	0.98	0.72
Il-Fawwara 2	1361	0.61	0.49	0.33	1.00	1.04
Il-Fawwara 3	1310	0.75	0.44	0.29	0.87	0.77
Il-Majjistral 1	3304	0.41	0.35	0.27	0.98	0.51
Il-Majjistral 2	2654	0.53	0.42	0.31	1.07	0.81
Il-Majjistral 3	1888	0.74	0.55	0.41	1.00	0.88
Il-Majjistral 4	2779	0.65	0.44	0.33	0.97	0.66
Il-Majjistral 5	2267	0.58	0.42	0.31	0.83	0.54
Il-Majjistral 6	546	0.70	0.34	0.23	0.95	1.47
Pembroke 1	753	0.54	0.57	0.45	0.70	0.72
Pembroke 2	3222	0.90	0.42	0.32	0.86	0.40
Pembroke 3	1046	0.55	0.54	0.44	0.81	0.77
Popeye 1	1845	0.59	0.34	0.24	1.05	0.80
Popeye 2	1502	0.55	0.32	0.25	1.03	0.90
Popeye 3	747	0.52	0.34	0.22	0.81	0.72
Selmun 1	2149	0.57	0.41	0.31	0.91	0.68
Selmun 2	1655	0.82	0.43	0.31	0.67	0.41
Selmun 3	1326	0.77	0.39	0.28	0.87	0.77
Wardija 1	1182	0.49	0.42	0.32	0.67	0.59
Wardija 2	896	0.73	0.46	0.34	0.79	0.84
Wardija 3	2174	0.31	0.33	0.23	0.70	0.28

A1.3. Photogrammetry accuracy by site

			Orthomosaic						DSM	
Site	Number of images	Flying altitude (m)	GSD (cm/pixel)	Reprojection error (pix)	GCP total RMSE (cm)	GCP X error (cm)	GCP Y error (cm)	GCP Z error (cm)	GSD (cm/pixel)	Point density (points/cm²)
Bahrija 1	380	52.9	1.13	0.765	3.3	1.96	1.67	2.13	2.25	0.197
Bahrija 2	325	50.1	1.07	0.729	7.8	1.77	1.39	7.49	2.15	0.217
Bahrija 3	326	53.8	1.16	0.917	6.5	1.82	3.08	5.48	2.32	0.186
Il-Fawwara 1	325	51.2	1.09	1.09	3	1.74	1.28	2.15	2.18	0.211
Il-Fawwara 2	349	53.4	1.1	1.02	7.3	2.7	6.32	2.61	2.2	0.207
Il-Fawwara 3	307	48.3	1.03	1.09	2.7	1.36	0.65	2.26	2.07	0.234
Il-Majjistral 1	328	36.8	0.8	0.812	1.3	0.52	0.71	0.97	1.59	0.394
Il-Majjistral 2	323	54.6	1.18	0.596	3.2	1.11	0.91	2.83	2.37	0.178
Il-Majjistral 3	343	62.7	1.33	0.665	6.8	4.28	5.03	1.73	2.67	0.141
Il-Majjistral 4	406	59.4	1.28	0.881	2.7	1.86	1.8	0.8	2.56	0.152
Il-Majjistral 5	325	44.1	0.94	0.8	5.9	1.26	0.87	5.73	1.88	0.281
Il-Majjistral 6	324	57.2	1.23	0.77	3.3	1	1.65	2.68	2.46	0.166
Pembroke 1	323	44.9	0.98	0.909	1.9	0.64	0.44	1.8	1.96	0.259
Pembroke 2	323	49.3	1.06	0.824	1.8	0.68	0.95	1.33	2.12	0.223
Pembroke 3	325	53.3	1.11	0.624	5.8	0.78	1.02	5.68	2.22	0.202
Popeye 1	325	49.1	1.04	0.757	3.8	1.62	0.74	3.33	2.08	0.23
Popeye 2	325	54.6	1.15	0.834	2.7	1.15	1.93	1.44	2.3	0.189
Popeye 3	324	53.3	1.14	0.89	4.2	3.08	2.68	0.76	2.27	0.194
Selmun 1	325	52.4	1.12	0.763	3.1	1.17	2.46	1.5	2.24	0.199
Selmun 2	323	46.2	0.98	0.806	3.1	0.53	0.96	2.94	1.97	0.258
Selmun 3	325	50.1	1.08	0.8	4.9	1.83	1.94	4.15	2.15	0.216
Wardija 1	325	45.6	0.97	0.878	2.4	1.47	1.5	1.2	1.95	0.263
Wardija 2	325	53	1.13	0.958	4.7	1.53	2.95	3.25	2.26	0.195
Wardija 3	325	34.9	0.77	0.782	2.9	0.89	1.4	2.35	1.54	0.421

A1.4. Complete data set by site

A1.4.1. Location, dependent variable, land covers set (LC, in m²) and shrubs metrics set (RS)

AOI	site	pellets_counts	LC1	LC2	LC3	LC4	LC5	RS1	RS2	RS4	RS5	RS6
Bahrija	bahri1	2	824	1096	3016	897	4165	1303	6884	38.35	0.98	104.7
Bahrija	bahri2	0	1274	864	2404	365	5090	571	6387	37.02	0.78	96.9
Bahrija	bahri3	82	1997	1514	2017	190	4256	295	6443	35.24	0.80	140.8
Il-Fawwara	faw1	10	2106	1068	1542	1514	3625	2139	7078	39.88	0.98	72.3
Il-Fawwara	faw2	39	1713	1746	2890	834	2371	1361	6125	48.87	1.00	104.4
Il-Fawwara	faw3	52	2652	1287	1813	980	3251	1310	7481	43.84	0.87	77.3
Il-Majjistral	majj1	120	1293	1451	2035	1346	3872	3304	4073	35.15	0.98	50.5
Il-Majjistral	majj2	101	1630	2045	1848	1414	3060	2654	5327	41.91	1.07	81.0
Il-Majjistral	majj3	100	1226	1837	2158	1395	3087	1888	7389	55.34	1.00	88.0
Il-Majjistral	majj4	147	1462	2145	743	1799	3848	2779	6474	44.35	0.97	65.9
Il-Majjistral	majj5	36	2412	2409	641	1307	3228	2267	5765	41.88	0.83	54.3
Il-Majjistral	majj6	199	2668	2047	1541	380	3360	546	6956	34.09	0.95	146.9
Pembroke	pem1	0	1739	1421	4141	409	2251	753	5434	57.26	0.70	71.7
Pembroke	pem2	0	1647	1271	2742	2909	1428	3222	9028	42.21	0.86	40.0
Pembroke	pem3	0	2677	1085	2577	572	3059	1046	5465	54.19	0.81	77.3
Popeye	pop1	195	1975	1439	1954	1091	3538	1845	5913	33.87	1.05	79.6
Popeye	pop2	7	1885	1862	1127	820	4303	1502	5462	32.37	1.03	90.4
Popeye	pop3	219	3401	1714	529	389	3963	747	5209	33.90	0.81	71.8
Selmun	selm1	28	1271	1171	3390	1230	2840	2149	5724	40.97	0.91	68.1
Selmun	selm2	237	573	1105	948	1350	6020	1655	8156	43.26	0.67	40.7
Selmun	selm3	358	655	1151	1647	1383	5160	1326	7700	39.00	0.87	76.6
Wardija	ward1	80	1538	777	1128	585	5955	1182	4948	42.35	0.67	59.2
Wardija	ward2	34	1089	375	4368	654	1680	896	7300	46.46	0.79	83.6
Wardija	ward3	154	1355	1320	1568	684	5061	2174	3146	32.83	0.70	27.6

A1.4.2. Anthropogenic variables (ANT) in m²

roads_100	roads_200	roads_500	rubbles_100	rubbles_200	rubbles_500	paths_100	paths_200	paths_500	agri_100	agri_200	agri_500	build_100	build_200	build_500
2088.3	3815.3	13822.9	1458.2	4869.6	22144.3	1286.9	3831.6	23797.6	15433.3	76663.0	289937.0	81.2	293.6	14342.9
539.2	8139.2	41276.7	896.0	4175.2	30053.0	2086.7	4852.7	22597.4	8174.4	37438.0	353923.1	910.1	9936.7	58987.4
1122.3	3272.7	15759.2	701.1	3624.5	23907.8	1764.2	4207.9	21906.5	2338.4	23195.8	289851.6	93.9	918.7	9820.9
0.0	2338.2	12392.0	2820.1	7344.0	25748.9	2696.1	8063.5	41931.3	16199.0	41744.4	185925.3	223.1	585.4	5398.9
590.6	3086.0	9118.7	2274.2	8004.5	30575.5	2521.3	4968.5	29792.9	10628.9	46843.9	272906.6	421.2	1770.9	6082.2
0.0	361.6	11209.8	2739.8	8256.3	38661.3	1078.9	5993.4	29810.4	12508.5	44618.1	343326.2	262.6	1852.6	7326.4
700.5	1005.8	15248.8	862.7	3438.0	21175.7	1989.0	5982.4	33748.5	443.0	11084.7	105830.4	86.4	1047.6	22056.0
0.0	0.0	2182.5	866.6	2122.8	11405.7	3183.8	7504.5	31534.1	7107.5	23091.3	35401.9	91.8	192.8	4428.8
1631.0	3567.0	9530.6	3906.9	8434.9	36287.4	2627.4	5025.4	23525.7	22485.1	68382.5	311695.4	179.4	381.9	9114.0
0.0	0.0	2057.9	1673.7	6992.0	31236.3	4358.8	8196.6	20629.6	8851.0	36986.5	178393.1	34.1	295.6	1076.6
0.0	0.0	0.0	1698.8	3569.9	8824.0	1347.7	5095.9	18864.3	606.2	806.9	13464.3	0.0	8.4	275.3
78.3	1693.7	9227.2	688.9	2747.5	18783.7	656.8	2016.4	12969.4	4352.7	26105.2	192642.8	29.2	902.8	11186.5
4617.8	19057.3	88035.5	335.7	811.1	3764.7	1124.7	2981.1	7558.6	0.0	0.0	11132.0	3722.7	22666.6	127675.9
4670.5	16684.0	86782.9	261.4	490.2	2929.7	1940.7	3483.0	10743.2	0.0	0.0	7159.4	4483.6	26716.3	121500.1
2441.2	7986.9	57621.8	624.7	1360.5	3737.7	343.5	2154.3	19182.8	0.0	0.0	917.1	854.4	7531.3	85734.1
0.0	496.7	7878.0	492.2	1247.1	12674.5	681.4	2231.4	16444.9	10269.3	15149.1	128406.7	13.4	2752.8	13860.3
0.0	0.0	13728.7	35.4	2741.9	24583.1	313.9	5810.0	21224.3	430.0	15983.2	182065.4	0.0	631.8	24454.2
61.5	2435.2	7558.0	225.3	923.8	13312.6	995.2	2644.0	12799.8	616.2	12597.6	101160.3	85.9	1099.8	7298.6
1931.0	9620.6	35305.8	1847.4	8838.6	53694.4	1053.7	2119.8	11747.6	6295.2	62424.1	360117.8	1176.7	7846.8	31701.4
460.7	6129.1	36326.5	923.8	5006.1	40226.8	2532.2	4659.2	11648.2	6.6	7118.7	155854.3	1621.3	13969.7	96629.4
0.0	97.1	17630.1	1619.3	9854.1	36298.8	1912.0	5082.1	13152.0	0.0	25509.4	128184.4	41.4	1025.8	54645.6
36.6	3820.3	15506.9	930.8	6252.4	36677.6	3321.7	7001.4	23648.3	457.5	33030.9	295570.2	107.2	1593.6	15708.7
1481.4	4074.3	16306.6	3520.4	10486.6	37734.5	2244.0	5871.4	23147.0	11765.0	49109.1	238131.5	2311.2	3113.2	18180.3
0.0	72.0	20388.2	2152.9	7788.8	42439.9	3070.4	6552.5	24724.0	7669.5	48540.5	336638.7	1346.9	4392.6	23913.7

Appendix 2 - Detailed methodology

A.2.1. Detailed photogrammetric process

The imagery captured by UAS was processed using Agisoft Metashape 1.8 software on an processing station. The Coordinate Reference System (CRS) of the Maltese CORS network is the European Datum 1950, projected in UTM 33N (EPSG: 23033). The CRS of DJI drones is WGS84 (EPSG: 4326), therefore, the images had their reference CRS converted using the transformation from WGS84 to ED50 for Malta (EPSG1144, x: -107, y: -88, z: -149). Low quality images were removed (<0.7) and the photos were aligned (HIGH quality, SOURCE pair preselection, key points limit: 40k, tie points limit: 10k, NO adaptive camera fitting). After the alignment, GCP were imported as markers and manually identified in 2 images per GCP before using the automatic detection feature to validate the remaining markers. Positioning information of the images were cleared and their alignment was performed again using only the markers' information. The camera locations were optimized before adjusting manually the bounding box of each SPC and computing DPCs (Quality: High, Depth Map filtering: Mild). The DPCs were used to produce orthomosaics and DSMs for each site (see section 4.2.1).

A.2.2 Detailed methodology of ground filtering approaches

Five methods of DTM generation (Table 3.2) were assessed for the accuracy of their absolute ground elevation (i.e. DTM accuracy) and vegetation height accuracy (i.e. AGM accuracy). Agisoft + CSF exhibited the highest accuracy (lower mean error), while the LiDAR method presented the greatest precision (lower standard deviation). However, the LiDAR method demonstrated superior consistency across diverse topography and terrain features, specifically concerning shrub heights, despite its lower initial spatial resolution. The Agisoft_CSF method demanded iterative parameter adjustments for each site across two different software programs, and this parameters selection could be the focus of an entire dedicated study (Baily et al. 2022). Iterative adjustment can potentially increase errors as attaining optimal settings is challenging, in addition to known issues over specific types of terrain such as terraced slopes (Pilj et al. 2020). Conversely, the LiDAR method provides more continuous ground points, because it can partially penetrate vegetation with a more homogeneous acquisition method. It reduces the variability of errors during DTM generation but introduces a somewhat constant offset, possibly due to the difficulty of LiDAR beams to penetrate denser low vegetation, classifying instead its surface as ground. Original LiDAR point clouds from the PA were clipped for each site and imported into CloudCompare 2.13, then filtered to retain points from the 'ground' class only. DTMs were created from the ground points by generating a mesh using a 2.5D Delaunay triangulation (best fitting plane), sampling points on the mesh (1000 pts/m²), and rasterizing the sampled points with a cell size of 0.15m, aligning with the original LiDAR data resolution (MEPA 2013). AGMs were computed for each site and negative values were set to 0. A robust linear regression model developed for the occasion was applied on above-ground features only (elevation > 0) to correct the relative elevation offsets introduced by the LiDAR DTMs.

A.2.3. Detailed methodology of classification post-processing

Polygons containing a maximum height of less than 0.25m or with an average area of less than 0.05m² were reclassified as 'low shrubs', while polygons containing a maximum height of more than 3m were reclassified into a new 'trees' category. The remaining class of 'tall shrubs' has been renamed 'relevant shrubs', corresponding to shrubs with a height of between 0.25m and 3m and a minimum surface area of 0.05m², which are hypothesized here to widely encompass the size range of woody vegetation likely to offer refuge to the rabbit (see section 2.2.2). Surface areas of the updated classes were computed from the reclassified rasters in m², and the 'relevant shrubs' polygons were used to resample the AGM layer for maximum and average individual shrub height. They were eventually converted to points to perform Average Nearest Neighbour (ANN) analysis where the distance between each shrub and its nearest neighbour was computed and averaged over each site (Observed average distance between shrubs). The expected average distance between shrubs is computed based on the hypothetical random distribution of the number of features over the surface area of a site (ESRI n.d. (b)). It is eventually divided by the observed average distance to produce the Nearest Neighbour Index, which gives a measure of dispersion relatively to a random distribution. Values below one indicates a clustered distribution, values around one indicates a random distribution while values above one indicates a dispersed distribution.

Appendix 3. Complete R code used for data analysis

```
library(readxl)
library(plyr)
library(dplyr)
library(DHARMa)
library(glmmTMB)
library(AICcmodavg)
library(moments)
library(MASS)
require(pscl)
library(overdisp)
library(ggplot2)
library(fitdistrplus)
library(AER)
library(lme4)
library(lmtest)
library(broom.mixed)
library(easyCODA)
library(compositions)
library(sjPlot)
library(tidyr)
library(countreg)
library(topmodels)
library(patchwork)

setwd("B:\\..WORKAREA\\Etudes\\Master\\RESEARCH\\Results")

#=========================================================================
#----------------------
# 0. DATA PREPARATION
#----------------------
#Load DF
data_site <- read_xlsx("JT_MSc_Rawdata_Dec23.xlsx", sheet = "LC_base2")
```

```
#Remove Pembroke sites from dataset
data_site<- data_site[data_site$AOI != "Pembroke", ]

#Standardize explanatory variables and store mean & SD for unstandardization
standardize_data <- function(data) {
  means <- apply(data, 2, mean, na.rm = TRUE)
  sds <- apply(data, 2, sd, na.rm = TRUE)
  standardized_data <- scale(data)
  return(list(standardized_data = standardized_data, means = means, sds = sds))
}

result_standardize <- standardize_data(data_site[,c(10:29)])
data_site[,c(10:29)]<-result_standardize$standardized_data

#-------------------
# Compositional data preparation
#-------------------
# create compositional explanatory set
df_LC <- data_site[, c("LC1", "LC2", "LC3", "LC4", "LC5")]

# Log-ratio selection with easyCODA
(step_LC2 <- STEP(df_LC))
LR2 <- step_LC2$logratios

#Create the result table
logratio_names <- step_LC2$names
R2_values <- step_LC2$R2max
pro_cor_values <- step_LC2$pro.cor
total_variance <- step_LC2$totvar
# Create a data frame
result_table <- data.frame(
  Logratio = logratio_names,
```

```
  R2 = R2_values,
  ProCor = pro_cor_values
)
result_table$TotalVariance <- total_variance
result_table

#Append selected LRs to data_site
data_site <- cbind(data_site, LR2)

#----------------
# Descriptive plots
#----------------
# LC dataset
data_site_sorted <- data_site %>% arrange(pellets_counts)
sorted_sitenames <- data_site_sorted$site
df_LC_sorted <- data_site_sorted[, c(5:9)]
long_data <- df_LC_sorted %>%
  mutate(site = factor(sorted_sitenames, levels = sorted_sitenames)) %>%
  pivot_longer(cols = -site, names_to = "Land_cover_type", values_to =
"Value") %>%
  group_by(site) %>%
  mutate(Proportion = Value / sum(Value) * 100)

# Define the names and colors for the land cover types
land_cover_names <- c("LC1" = "Rock (LC1)", "LC2" = "Bare soil (LC2)",
                      "LC3" = "Herbaceous (LC3)", "LC4" = "Relevant shrubs
(LC4)",
                      "LC5" = "Low shrubs (LC5)")
land_cover_colors <- c("#CCCCCC", "#FF9900", "#99CC33", "#006600",
"#336666")

# Create the plot
ggplot(long_data, aes(x = site, y = Proportion, fill = Land_cover_type)) +
  geom_bar(stat = "identity") +
```

```
  scale_fill_manual(values = land_cover_colors, labels = land_cover_names)
+
  labs(x = "Site", y = "Proportion of site surface (% of 1ha)", fill = "Land
cover type") +
  theme(axis.text.x = element_text(angle = 45, hjust = 1, size = 12),
        axis.text.y = element_text(size = 12),
        axis.title.x = element_text(margin = margin(t = 10)),
        axis.title.y = element_text(margin = margin(r = 10))) +
  coord_flip() +
  guides(fill=guide_legend(title="Land cover type", reverse=TRUE))


#=============================================================================
# 1.DISTRIBUTIONS OF VARIABLES
#=============================================================================
#---------------------------
# 1.1 Dependent variable
#---------------------------


#Extract the raw DV
pel_site <- data_site$pellets_counts


#Outliers
par(mfrow = c(1, 1))
boxplot(pel_site, ylab="Pellets/site")
# result = one outlier (selm3)
# not too bad, and real data = keep it


#histograms
hist(pel_site, breaks = "fd", xlab = "Total pellets per site", freq = TRUE,
main = "")
# result = slightly right-skewed with higher 1st bar


#Descriptive statistics
(mean_pel <- mean(pel_site))
(sd_pel <- sd(pel_site))
```

```
(var_pel <- var(pel_site))
(kurt_pel <- kurtosis(pel_site))
(skew_pel <- skewness(pel_site))
(min(pel_site))
(max(pel_site))

# Variance (9005.39) is larger than the mean (104.76)
# SD (94.9) is lower than the mean
# skewness: 0.97 - indicates right-skewed data
# Kurtosis: 3.4 - slightly leptokurtic (slightly more peaky than normal)

#Check for zero-inflation
zero_count<-NROW(data_site_RS[which(data_site_RS$pellets==0),])
zero_prop<-(zero_count/NROW(data_site_RS))*100

#Low proportion of sites with zeros pellet counted = 4.8%

#=============================================================================
#A. SHRUB METRICS
#=============================================================================
#-----------------
# Model structure
#-----------------
#  Poisson regression
mod.p<- glm(pellets_counts ~ RS1+RS2+RS4+RS5+RS6,
            family="poisson",
            data=data_site)

# Test for dispersion (over 1 = overdispersed, under 1 = underdispersed.
NULL = all random effects)
sim_resids_mod.p <- simulateResiduals(mod.p, re.form = NULL, n=300)
plot(sim_resids_mod.p)
testDispersion(sim_resids_mod.p)
```

```
# results = Poisson GLM fails to generate scaled residuals approximating uniform distribution

# results = significant overdispersion of data (61.9, significant p-value, red bar far away)


# Test for zero inflation (over 1 = more zeros in observed data than expected by model (zero-inflated))

testZeroInflation(sim_resids_mod.p)

#Results = significant zero inflation


summary(mod.p)

AICc(mod.p)


# result = Poisson GLM is zero-inflated & overdispersed

# Consider models accounting for overdispersion and zero-inflation


#-------------

# Mean-Variance plot

#-------------

# Calculate predicted values and squared residuals

predicted_p <- predict(mod.p, type = "response")

residuals_p <- data_site$pellets_counts - predicted_p

squared_residuals_p <- residuals_p^2


# Create bins for observed values

data_site$observed_bins <- cut(data_site$pellets_counts, breaks = 5)


# Calculate average observed values and average squared residuals for each bin

average_data <- aggregate(cbind(mean_observed = data_site$pellets_counts, mean_squared_residual_p = squared_residuals_p),

                          by = list(data_site$observed_bins),

                          FUN = mean, na.rm = TRUE)


# Plotting with legends for the linear and quadratic trends
```

```
ggplot(average_data, aes(x = mean_observed, y = mean_squared_residual_p)) +
  geom_point() +
  geom_smooth(aes(color = "Linear Trend"), method = "lm", formula = y ~ x)
+
  geom_smooth(aes(color = "Quadratic Trend"), method = "lm", formula = y ~
poly(x, 2, raw = TRUE)) +
  scale_color_manual(values = c("Linear Trend" = "blue", "Quadratic Trend"
= "red")) +
  ggtitle("Poisson model: Observed vs Squared Residuals") +
  theme(legend.title = element_blank())


# Results = mean-variance seems to be quadratic (NBIN2)


#------------
# NEGATIVE BINOMIAL
#------------


# NBINOM2 = quadratic relationship between variance & mean
# (variance = μ(1 + μ/k) with mean=u and k=dispersion parameter)
mod.nb2<-glmmTMB(pellets_counts ~ RS1+RS2+RS4+RS5+RS6,
                 family="nbinom2",
                 data=data_site)
sim_resids_mod.nb2 <- simulateResiduals(mod.nb2, re.form = NULL, n = 500)
plot(sim_resids_mod.nb2)
testZeroInflation(sim_resids_mod.nb2)
testDispersion(sim_resids_mod.nb2)
summary(mod.nb2)
AICc(mod.nb2)


#NBINOM1 = linear relationship between variance & mean
# variance = μ(1 + k), with k > 0;
mod.nb1<-glmmTMB(pellets_counts ~  RS1+RS2+RS4+RS5+RS6,
                 family="nbinom1",
                 data=data_site)
sim_resids_mod.nb1 <- simulateResiduals(mod.nb1, re.form = NULL, n = 300)
```

```
plot(sim_resids_mod.nb1)
test2 <- testZeroInflation(sim_resids_mod.nb1)
testDispersion(sim_resids_mod.nb1)
summary(mod.nb1)
AICc(mod.nb1)

#--------------
# variable selection
#--------------
RS <- c("RS1", "RS2", "RS4", "RS5", "RS6")

full <- data.frame(full="pellets_counts ~ RS1+RS2+RS4+RS5+RS6")
Dot<-data.frame(full="pellets_counts ~ 1")

single<-expand.grid(RS)
double<-as.data.frame(t(combn(RS,2)))
names(double)<-c("Var1","Var2")

triple<-as.data.frame(t(combn(RS,3)))
names(triple)<-c("Var1","Var2", "Var3")

quad<-as.data.frame(t(combn(RS,4)))
names(quad)<-c("Var1","Var2", "Var3", "Var4")

single$full<-paste("pellets_counts ~ ", single$Var1)
double$full<-paste(double$Var1, double$Var2, sep="+")
double$full<-paste("pellets_counts ~ ", double$full)

triple$full<-paste(triple$Var1, triple$Var2, triple$Var3, sep="+")
triple$full<-paste("pellets_counts ~ ", triple$full)

names(quad)<-c("Var1","Var2", "Var3", "Var4")
quad$full<-paste(quad$Var1, quad$Var2, quad$Var3, quad$Var4, sep="+")
```

```
quad$full<-paste("pellets_counts ~ ", quad$full)


combo<-data.frame(full=c(Dot$full,   full$full,   single$full,double$full,
triple$full, quad$full))


Stage3=list()

Stage3_formulas<-as.list(combo$full)

Stage3_formulas<-lapply(Stage3_formulas, as.formula)


for (i in seq_along(Stage3_formulas)) {


  Stage3[[i]]<-glmmTMB(Stage3_formulas[[i]],              family="nbinom2",
ziformula=~1, data=data_site)

  names(Stage3)[[i]]<-combo$full[[i]]

}


# Calculate AICc values to compare models

Stage3_AIC<-aictab(Stage3, second.ord=T)

Stage3_TopAIC<-Stage3_AIC[which(Stage3_AIC$AICcWt > 0.05),]


#------

# Alternative selection method: stepwise selection

#------

RS_fullmodel <- mod.RS_zinbin2

RS_nullmodel <- glmmTMB(pellets_counts ~  1,

                        family="nbinom2",

                        ziformula=~1,

                        data=data_site)


forward_model <- step(RS_nullmodel, direction="forward",

                      scope=list(lower=~1, upper=~RS1+RS2+RS4+RS5+RS6),

                      data=your_data)

backward_model <- step(RS_fullmodel, direction="backward")

both_model <- step(RS_fullmodel, direction="both")
```

```
summary(both_model)


# Same results


#----------
# No significant predictors in shrub metrics - null model has the best AIC, no predictors retained in stepwise
# Move on to land covers


#=========================================================================
# B. LAND COVERS
#=========================================================================
#-------------------------
# Model structure with STEP
#-------------------------
mod.p<- glm(pellets_counts ~ `LC1/LC3`+`LC3/LC5`+`LC4/LC5`+`LC2/LC5`,
            family="poisson",
            data=data_site)


#-------------
# Mean-Variance plot with STEP
#-------------
# Calculate predicted values and squared residuals
predicted_p <- predict(mod.p, type = "response")
residuals_p <- data_site$pellets_counts - predicted_p
squared_residuals_p <- residuals_p^2


# Create bins for observed values
data_site$observed_bins <- cut(data_site$pellets_counts, breaks = 5)


# Calculate average observed values and average squared residuals for each bin
average_data <- aggregate(cbind(mean_observed = data_site$pellets_counts, mean_squared_residual_p = squared_residuals_p),
                          by = list(data_site$observed_bins),
```

```
                          FUN = mean, na.rm = TRUE)

# Plotting Poisson binned observed VS residuals
ggplot(average_data, aes(x = mean_observed, y = mean_squared_residual_p)) +
  geom_point() +
  geom_smooth(aes(color = "Linear Trend"), method = "lm", formula = y ~ x)
+
  geom_smooth(aes(color = "Quadratic Trend"), method = "lm", formula = y ~
poly(x, 2, raw = TRUE)) +
  scale_color_manual(values = c("Linear Trend" = "blue", "Quadratic Trend"
= "red")) +
  ggtitle("Poisson model: Observed vs Squared Residuals") +
  theme(legend.title = element_blank())

# Results = mean-variance seems to be quadratic (NBIN2)

#----------------------------
# NEGATIVE BINOMIAL with STEP
#----------------------------

# NBINOM2
mod.nb2<-glmmTMB(pellets_counts ~ `LC1/LC3`+`LC3/LC5`+`LC4/LC5`+`LC2/LC5`,
                 family="nbinom2",
                 data=data_site)



sim_resids_mod.nb2 <- simulateResiduals(mod.nb2, re.form = NULL, n = 500)
plot(sim_resids_mod.nb2)
testZeroInflation(sim_resids_mod.nb2)
testDispersion(sim_resids_mod.nb2)
summary(mod.nb2)
AICc(mod.nb2)

# No significant issues with NB2
```

```
#--------------
# variable selection
#--------------
LC <- c("`LC1/LC3`","`LC3/LC5`", "`LC4/LC5`", "`LC2/LC5`")


full            <-            data.frame(full="pellets_counts            ~
`LC1/LC3`+`LC3/LC5`+`LC4/LC5`+`LC2/LC5`")
Dot<-data.frame(full="pellets_counts ~ 1")


single<-expand.grid(LC)
double<-as.data.frame(t(combn(LC,2)))
names(double)<-c("Var1","Var2")
triple<-as.data.frame(t(combn(LC,3)))
names(triple)<-c("Var1","Var2", "Var3")


single$full<-paste("pellets_counts ~ ", single$Var1)
double$full<-paste(double$Var1, double$Var2, sep="+")
double$full<-paste("pellets_counts ~ ", double$full)
triple$full<-paste(triple$Var1, triple$Var2, triple$Var3, sep="+")
triple$full<-paste("pellets_counts ~ ", triple$full)


combo<-data.frame(full=c(Dot$full,   full$full,   single$full,double$full,
triple$full))


Stage3=list()
Stage3_formulas<-as.list(combo$full)
Stage3_formulas<-lapply(Stage3_formulas, as.formula)


for (i in seq_along(Stage3_formulas)) {


  Stage3[[i]]<-glmmTMB(Stage3_formulas[[i]],              family="nbinom2",
data=data_site)
  names(Stage3)[[i]]<-combo$full[[i]]
}
```

```
# Calculate AICc values to compare models
Stage3_AIC<-aictab(Stage3, second.ord=T)
Stage3_TopAIC<-Stage3_AIC[which(Stage3_AIC$AICcWt > 0.05),]

#-------------------
# Stepwise selection
#----------------

LC_nullmodel <- glmmTMB(pellets_counts ~  1,
                            family="nbinom2",
                            data=data_site)

summary (LC_nullmodel)
tab_model(LC_nullmodel)

forward_model <- step(LC_nullmodel, direction="forward",
                        scope=list(lower=~1,
upper=~`LC1/LC3`+`LC3/LC5`+`LC4/LC5`+`LC2/LC5`),
                        data=data_site)
backward_model <- step(mod.nb2, direction="backward")
both_model <- step(mod.nb2, direction="both")
summary(LC_nullmodel)

# LC3/LC5 retained with weak support (low dAICc with null model)
# Similar outcome than full model
# Keep full model for interpretation and present single model for precision

tab_model(mod.nb2)

#-------------
# LC final model
#--------------
mod.nb2_final<-glmmTMB(pellets_counts ~ `LC3/LC5`,
                  family="nbinom2",
```

```
                  data=data_site)
sim_resids_mod.nb2_final <- simulateResiduals(mod.nb2_final, re.form = NULL,
n = 500)
plot(sim_resids_mod.nb2_final)
testZeroInflation(sim_resids_mod.nb2_final)
testDispersion(sim_resids_mod.nb2_final)
summary(mod.nb2_final)
AICc(mod.nb2_final)


tab_model(mod.nb2_final)
plot_model(mod.nb2_final, terms = "LC3/LC5", type = "pred", show.data=TRUE)


# Compare full and single model with LR test (nested models)
lmtest::lrtest(mod.nb2, mod.nb2_final)


#----------------------
# Sensitivity analysis with Corrected LC dataset
#---------------------
#Load DF
data_site_cor    <-    read_xlsx("JT_MSc_Rawdata_Dec23.xlsx",    sheet    =
"LC_corrected_sensitivity")


#Remove Pembroke
data_site_cor<- data_site_cor[data_site_cor$AOI != "Pembroke", ]


# create compositional set
df_LC3 <- data_site_cor[, c("LC1", "LC2", "LC3", "LC4", "LC5")]
(step_LC3 <- STEPR(df_LC, data_site_cor$pellets_counts, family = "poisson",
method=1))
LR3 <- step_LC3$logratios


#Append selected LRs to data_site
data_site_cor<- data_site_cor[-c(5:8)]
data_site_cor <- cbind(data_site_cor, LR3)
```

```
LC <- c("`LC3/LC5`","`LC1/LC2`", "`LC4/LC5`")

full <- data.frame(full="pellets_counts ~ `LC3/LC5`+`LC1/LC2`+`LC4/LC5`")
Dot<-data.frame(full="pellets_counts ~ 1")

single<-expand.grid(LC)
double<-as.data.frame(t(combn(LC,2)))
names(double)<-c("Var1","Var2")
triple<-as.data.frame(t(combn(LC,3)))
names(triple)<-c("Var1","Var2", "Var3")

single$full<-paste("pellets_counts ~ ", single$Var1)
double$full<-paste(double$Var1, double$Var2, sep="+")
double$full<-paste("pellets_counts ~ ", double$full)
triple$full<-paste(triple$Var1, triple$Var2, triple$Var3, sep="+")
triple$full<-paste("pellets_counts ~ ", triple$full)

combo<-data.frame(full=c(Dot$full,   full$full,   single$full,double$full,
triple$full))

Stage3=list()
Stage3_formulas<-as.list(combo$full)
Stage3_formulas<-lapply(Stage3_formulas, as.formula)

for (i in seq_along(Stage3_formulas)) {

  Stage3[[i]]<-glmmTMB(Stage3_formulas[[i]],              family="nbinom2",
data=data_site_cor)
  names(Stage3)[[i]]<-combo$full[[i]]
}

# Calculate AICc values to compare models
Stage3_AIC<-aictab(Stage3, second.ord=T)
Stage3_TopAIC<-Stage3_AIC[which(Stage3_AIC$AICcWt > 0.05),]
```

```
# LC3/LC5 still retained

mod.nb3_final<-glmmTMB(pellets_counts ~ `LC3/LC5`,
                       family="nbinom2",
                       data=data_site_cor)
sim_resids_mod.nb3_final <- simulateResiduals(mod.nb3_final, re.form = NULL,
n = 500)
plot(sim_resids_mod.nb3_final)
testZeroInflation(sim_resids_mod.nb3_final)
testDispersion(sim_resids_mod.nb3_final)
summary(mod.nb3_final)
AICc(mod.nb3_final)

# LC3/LC5 still retained
tab_model(mod.nb3_final)
plot_model(mod.nb3_final, terms = "LC3/LC4", type = "pred", show.data=TRUE)

#=========================================================================
# C. ANTHROPOGENIC
#=========================================================================
# Prepare variables
scales <- c(100, 200, 500)
variables <- c("roads_", "paths_", "rubbles_")
variable_names <- sort(outer(variables, scales, paste0))

# Create descriptive statistics for all scales
summary_df <- lapply(variable_names, function(var) {
  data_subset <- data_site_original[[var]]
  tibble(
    Variable = var,
    Mean = mean(data_subset, na.rm = TRUE),
    SD = sd(data_subset, na.rm = TRUE),
    Min = min(data_subset, na.rm = TRUE),
```

```
    Max = max(data_subset, na.rm = TRUE),
    Median = median(data_subset, na.rm = TRUE),
    Zero_Count = sum(data_subset == 0, na.rm = TRUE)
  )
}) %>% bind_rows()

#-----------------
# Model structure
#-----------------
#  Poisson regression
mod.p<- glm(pellets_counts ~ roads_500+paths_500+rubbles_500,
            family="poisson",
            data=data_site)

sim_resids_mod.p <- simulateResiduals(mod.p, re.form = NULL)
testDispersion(sim_resids_mod.p)
plot(sim_resids_mod.p)
testZeroInflation(sim_resids_mod.p)
summary(mod.p)
AICc(mod.p)

#-------------
# Mean-Variance plot
#-------------
# Calculate predicted values and squared residuals
predicted_p <- predict(mod.p, type = "response")
residuals_p <- data_site$pellets_counts - predicted_p
squared_residuals_p <- residuals_p^2

# Create bins for observed values
data_site$observed_bins <- cut(data_site$pellets_counts, breaks = 5)

# Calculate average observed values and average squared residuals for each
bin
```

```
average_data <- aggregate(cbind(mean_observed = data_site$pellets_counts,
mean_squared_residual_p = squared_residuals_p),
                          by = list(data_site$observed_bins),
                          FUN = mean, na.rm = TRUE)


# Plotting Poisson binned observed VS residuals
ggplot(average_data, aes(x = mean_observed, y = mean_squared_residual_p)) +
  geom_point() +
  geom_smooth(aes(color = "Linear Trend"), method = "lm", formula = y ~ x)
+
  geom_smooth(aes(color = "Quadratic Trend"), method = "lm", formula = y ~
poly(x, 2, raw = TRUE)) +
  scale_color_manual(values = c("Linear Trend" = "blue", "Quadratic Trend"
= "red")) +
  ggtitle("Poisson model: Observed vs Squared Residuals") +
  theme(legend.title = element_blank())


# Results = mean-variance seems to be quadratic (NBIN2)


#-----------
# NEGATIVE BINOMIAL
#-----------


# NBINOM2
mod.nb2<-glmmTMB(pellets_counts ~ roads_500+paths_500+rubbles_500,
                 family="nbinom2",
                 data=data_site)
sim_resids_mod.nb2 <- simulateResiduals(mod.nb2, re.form = NULL, n = 500)
plot(sim_resids_mod.nb2)
testZeroInflation(sim_resids_mod.nb2)
testDispersion(sim_resids_mod.nb2)
summary(mod.nb2)
AICc(mod.nb2)


#----------
```

```
# Variable selection
#----------

# Define categories
roads <- c("roads_100", "roads_200", "roads_500")
paths <- c("paths_100", "paths_200", "paths_500")
rubbles <- c("rubbles_100", "rubbles_200", "rubbles_500")
agri <- c("agri_100", "agri_200", "agri_500")
build <- c("build_100", "build_200", "build_500")
categories <- list(roads, paths, rubbles)

# Function to generate combinations for n categories
generate_combinations_n <- function(categories, n) {
  combos <- combn(categories, n, simplify = FALSE)
  model_formulas <- lapply(combos, function(combo) {
    cat_combos <- expand.grid(combo)
    formulas <- apply(cat_combos, 1, function(row) {
      paste("pellets_counts ~", paste(row, collapse = " + "))
    })
    return(formulas)
  })
  return(unlist(model_formulas))
}

# Generate model formulas for combinations of 3 predictors max.
all_formulas <- c()
for (n in 1:3) {
  all_formulas <- c(all_formulas, generate_combinations_n(categories, n))
}

# Add the NULL model
all_formulas <- c("pellets_counts ~ 1", all_formulas)
```

```
# Remove duplicates and standardize order
all_formulas <- unique(sapply(all_formulas, function(formula) {
  parts <- strsplit(formula, " \\+ ")[[1]]
  sorted_parts <- c(parts[1], sort(parts[-1]))
  paste(sorted_parts, collapse = " + ")
}))

# Convert to dataframe
model_formulas_df <- data.frame(full = all_formulas)
head(model_formulas_df)

# Compute models AIC, AIC weights and likelihood
Stage3=list()
Stage3_formulas<-as.list(model_formulas_df$full)
Stage3_formulas<-lapply(Stage3_formulas, as.formula)

for (i in seq_along(Stage3_formulas)) {

  Stage3[[i]]<-glmmTMB(Stage3_formulas[[i]],              family="nbinom1",
data=data_site)
  names(Stage3)[[i]]<-model_formulas_df$full[[i]]
}

# Calculate AICc values to compare models
Stage3_AIC<-aictab(Stage3, second.ord=T)
Stage3_TopAIC<-Stage3_AIC[which(Stage3_AIC$AICcWt > 0.05),]

# Best model is roads_200+paths_500 followed by same+rubbles_500
# 3 predictors might be overfitting for 21 observations (1 predictor / 10-
20 obs = 1 or 2 predictors max)



anthro_fullmodel <- glmmTMB(pellets_counts ~ raods_200 + paths_500,
```

```
                              family="nbinom2",
                              data=data_site)
sim_resids_anthro_fullmodel <- simulateResiduals(anthro_fullmodel, re.form
= NULL, n = 500)
plot(sim_resids_anthro_fullmodel)
testZeroInflation(sim_resids_anthro_fullmodel)
testDispersion(sim_resids_anthro_fullmodel)
summary(anthro_fullmodel)
AICc(anthro_fullmodel)


# Exponentiate model and and create predictive plots for each predictor
tab_model(anthro_fullmodel)
plot1 <- plot_model(anthro_fullmodel, type = "pred", terms = "roads_200",
show.data = TRUE) +
  ggtitle("")
plot2 <- plot_model(anthro_fullmodel, type = "pred", terms = "paths_500",
show.data = TRUE) +
  ggtitle("")
combined_plot <- plot1 / plot2
combined_plot
```