Assessing the habitat response of breeding farmland waders in Shetland

Anthony Waite

17 September 2020

# Abstract

To do.

# Acknowledgements

“When we try to pick out anything by itself, we find it hitched to everything else in the universe.” - John Muir

# Introduction

## Farmland waders

Oystercatcher, Lapwing, Curlew, Redshank and Snipe are birds within the order *Charadriiformes* and are commonly referred to as *farmland waders*, due to their usage of agricultural land to ground nest, often in *plain site*. This nesting strategy requires that a nesting wader continuously surveils their environment for predators. If the nest is threatened, the nesting birds will immediately take flight, relying on the fact that eggs within the nest are cryptically marked so as to avoid detection by predators. A short description of each farmland wader and it’s conservation status is summarised below:

* **Euraisian Oystercatcher** *Haematopus ostralegus* - primarily lives at the coast, where it also breeds and feeds on bivalves and gastropods. Has over the last 50 years started to breed in-land on farmland, where it’s primary food source is earthworms and insect larvae. Oystercatcher have strong mate and site fidelity, where a single nesting attempt is made per breeding season. They typically lay their eggs in April and unlike other farmland waders, they will sometimes nest close to man made habitat such as buildings or on top of drystone walls. Nesting on flat gravel roofs was first noted in Aberdeen in the 1960s, with these birds often feeding in urban parks and on lawns. Unlike other waders, adults can take food to their young rather than requiring the young to move to areas where they can feed themselves; a feature which enables rooftop nesting. The current IUCN Red List status for the Oystercatcher is *near threatened* (BirdLife International [2019](#ref-BirdLife_International2019-tc)).
* **Northern Lapwing** *Vanellus vanellus* - outside of the breeding season, they may overwinter in large flocks in-land as far south as northern Africa. Lapwing feed on insects and earthworms, and have a preference for feeding at night in order to avoid predators. As the Lapwing has a short bill, they can only forage for food near the surface of the soil. They will therefore typically nest in areas that are invertebrate rich. Lapwing often form loose colonies when nesting, that can together drive away predators. Their chicks typically hatch in early May, with replacement clutches typically in June if the first hatch failed. The current IUCN Red List status for the Lapwing is *vulnerable* (BirdLife International [2017](#ref-BirdLife_International2017-vv)).
* **Eurasian Curlew** *Numenius arquata* - Scotland’s largest waders generally spend their winters near the coast and return to their Shetland breeding territories to hatch their chicks in early June. This later breeding period than Oystercatchers and Lapwings is associated with a requirement for the nest to be in tall vegetation (c.15cm tall or more). The Curlew can be migratory, often over wintering in Africa or Southern Europe. Though in the milder climate of Scotland it can be resident all year round. The current IUCN Red List status for the Curlew is *vulnerable* (BirdLife International [2015b](#ref-BirdLife_International2015-or)).
* **Common Redshank** *Tringa totanus* - perhaps the least common of the farmland waders. They spend the winter at the coast, particularly on mudflats and estuaries and nest most abundantly on wetland, especially coastal saltmarsh and machair. Pools of shallow standing water are particularly attractive to nesting Redshank, as they feed on the small invertebrates that they support. Their chicks hatch in late May. They are a migratory species, over wintering on Atlantic coasts south of Ireland and Great Britain. The current IUCN Red List status for the Redshank is *least concern* (BirdLife International [2015c](#ref-BirdLife_International2015-yo))
* **Common Snipe** *Gallinago gallinago* - Snipe nest in dense cover of rushes and sedges within wetland habitats, as well as in wet scrapes and bogs in moorland habitats. They feed on invertabrates in areas of shallow standing water. They reside in farmland wetlands throughout the year. Snipe are often difficult to observe and often are only detectable by listening out for the *drumming* that results from their display flights. The current IUCN Red List status for the Snipe is *least concern* (BirdLife International [2015a](#ref-BirdLife_International2015-gf))

## Declines in wader populations

Wader populations are declining across the world (Butchart et al. [2010](#ref-Butchart2010-yf)), and farmland waders are one of the highest wildlife conservation priorities on Scottish farmland. The loss and degradation of natural and semi-natural breeding habitats through changing land use, particularly agricultural intensification, changes in cropping and grazing management is believed to have caused significant declines in farmland wader populations (Thorup [2006](#ref-Thorup2006-zd)). It is generally accepted that the main cause of this reduction lies in *low breeding productivity* which in-turn is most likely due to agricultural intensification (Wilson, Ausden, and Milsom [2004](#ref-Wilson2004-mr)). A recent paper (Bell and Calladine [2017](#ref-Bell2017-iy)) reported results of a 25 year study (1990 to 2015) on a typical Scottish farm north of Stirling (a survey area of 7.5km). The authors explain how over this time Oystercatcher have declined by -95%, Lapwing by -88%, Curlew by -67% and Redshank by -87%. The main reason cited for the declines were changes in agricultural practices relating to crop cycles and grazing. This declining trend is reinforced by the annual BTO Breeding Bird Survey for 2017 (Harris et al. [2018](#ref-Harris2018-mt)). Over a similar time scale as the Stirling study, the UK wide trend between 1995 and 2016 has seen declines for Oystercatcher of -23%, Redshank -41%, Lapwing -42% and Curlew -48%. Snipe were the only species to increase over this time period, by +26%.

## Breeding wader abundance response to farmland habitat

The suitability of farmland habitat for breeding waders is likely to be influenced by the historic conversion of the land from more heterogeneous semi-natural grassland to agriculturally improved grassland through soil drainage, liming, fertilization and re-seeding. These practices can result in a structurally uniform sward with reduced invertebrate and earthworm density (Vickery et al. [2001](#ref-Vickery2001-qf)), though improvement through fertilization and liming (McCallum et al. [2018](#ref-McCallum2018-gx)) may increase the availability of some soil invertebrates. Lowland wet unimproved grassland can provide suitable nesting sites and are a good source of soil invertebrates (Smart et al. [2008](#ref-Smart2008-pk)), as well as foraging sites for birds breeding on nearby unenclosed moorland (Brown et al. [2015](#ref-Brown2015-ay)), particularly if fields have high densities of soil invertebrates.

Agriculture is generally thought to depress biodiversity but in some cases it can be beneficial. It has been postulated that there is a beneficial relationship between agricultural *improvement* of semi-natural land and species alpha diversity (Jóhannesdóttir et al. [2019](#ref-Johannesdottir2019-uh)), but that this relationship is unimodal. Agricultural practices such as liming, spreading animal manure on fields and hay cropping may increase habitat and resource heterogeneity, together with enhancing soil health. This leads to increases in alpha diversity, but beyond a critical level of intensity, alpha diversity quickly falls away due to the habitat homogenisation associated with more intensive and extensive cultivation. For example, a study on the interacting effects of agriculture and landscape on breeding wader populations in Iceland (Jóhannesdóttir et al. [2019](#ref-Johannesdottir2019-uh)) concludes that low intensity agriculture can have a positive effect on breeding wader populations in upland areas where food resources are scarce, but in coastal lowlands wader numbers were seen to decline with increasing cultivation of the land. Further, the proximity and extent of lowland wetlands within Iceland were a particularly important driver of farmland breeding wader abundance. A study to examine how environmental covariates across the UK are associated with breeding Curlew (Franks et al. [2017](#ref-Franks2017-co)) found that arable farming, generalist predators and woodland cover were negatively associated with Curlew abundance. The same study showed that Curlew abundance was positively associated with the extent of protected area coverage, gamebird numbers, cooler seasonal temperatures and higher summer rainfall.

A key requirement for breeding waders on farmland is the type of habitat and sward height. Table 1 summarises typical habitat and sward height that farmland waders typically prefer for their nesting site.

Table 1: Habitat preferences for breeding wader nest sites.

|  |  |  |  |
| --- | --- | --- | --- |
| Wader | Breeding habitat | Sward height | Breeding season |
| Oystercatcher | Grassland or arable fields | Short | April - July |
| Lapwing | Grassland, spring crops, fallow land | Bare/short | April - late July |
| Curlew | Heathland, grassland and bog | Long | May - late July |
| Redshank | Wet grassland, wet scrapes | Long | May - late July |
| Snipe | Wetland, standing water | Long | April - July |

With regards to the categorisation of *sward height*:

* **bare** - ploughed land with no emergent vegetation. This is typical of spring sowed crops
* **short** - managed (improved) grassland for grazing, or for mowing to make hayage or silage. Sward is typically less than 5cm tall.
* **long** - unimproved grassland, fallow pasture, heathland, marsh or wetland. Sward is typically 15cm or more.

## Breeding wader abundance response to landscape composition and scale

The size and composition of a landscape is key to sustaining a long term population of a bird species, especially one that has experienced a recent decline (Macdonald [2019](#ref-Macdonald2019-jt)). If a suitable landscape exists at the necessary scale, it will provide security against localised adverse effects such as starvation, predation, inbreeding and bad weather. This in-turn will allow populations, whose number is above a certain threshold, to survive major short-term fluctuations, in abundance (Melbourne and Hastings [2008](#ref-Melbourne2008-vg)). Understanding how abundance of breeding waders respond to the scale and composition of suitable habitat could be a major factor in helping to conserve them.

## The Shetland archipelago

Shetland is a subarctic archipelago located in the North Atlantic between Scotland, Norway and the Faroe Islands. The total area of the islands are some 1,466 km and they experience an oceanic temperate maritime climate.Due to the relatively harsh climate, short growing season and poor soil quality, Shetland has relatively *low agricultural intensification* when compared to the rest of the UK. The traditional crofting approach to farming has created farmland characterised by a mosaic of low intensity cropping and grazing. These factors have helped to create a hotspot for farmland breeding waders; the most recent estimates of the Shetland breeding population are from surveys carried out in the late 1990s: 2,300 pairs of Curlew, 1,740 pairs of Lapwing, 3,350 pairs of Oystercatcher, 1,170 pairs of Redshank and 3,450 pairs of Snipe (*The Birds of Shetland* [2004](#ref-Harvey2004-pm)).

## Study objective and hypotheses

The objective of this study is to quantify breeding wader abundance response to various environmental covariates, across the archipelago of Shetland. If agricultural practices and landscape scape changes are a significantly associated with the decline in farmland wader populations, Shetland’s lower intensity farming and remoteness, should generate a different response in breeding wader population trends than that seen on the UK mainland. The study will investigate how habitat, topography and landscape pattern are associated with breeding wader abundance on Shetland. Whilst there have been numerous UK studies on breeding wader abundance and habitat associations, none have examined associations across Shetland using a long term dataset. For this study, data from the Shetland Breeding Bird Survey (BBS) from 2002 to 2019 was used (ref), together with a number of spatial datasets that represent environmental covariates.

A number of hypotheses were formulated *a priori* as outlined in Table 2. Here predictions are made as to how the abundance per unit area (density) of breeding waders is associated with a particular covariate. Through the methods of the study, each individual hypothesis was tested for statistical significance.

Table 2: Hypothesis for associations of habitat scale and composition for breeding wader density.

|  |  |  |
| --- | --- | --- |
| ID | Hypothesis | Wader Density |
| **H1** | Farmland that is managed as improved grassland for silage has a higher pH, possibly due to liming and fertilisation, which in turn is associated with higher earthworm density (McCallum et al. [2016](#ref-McCallum2016-jt)) | Positive |
| **H2** | Earthworm density is negatively associated with high organic carbon content within the topsoil, due to increased soil acidity (McCallum et al. [2016](#ref-McCallum2016-jt)). | Negative |
| **H3** | Earthworm density is positively associated with higher topsoil pH (between 5.0-7.4), as common earthworm species are acid intolerant. (McCallum et al. [2016](#ref-McCallum2016-jt)) | Positive |
| **H4** | Heathland is associated with insects such as Crane fly and their larvae (*Tipulidae spp.*), and are a major food source for wader chicks | Positive |
| **H5** | Semi-natural grasslands are associated with both earthworms and invertebrates, and are a major food source for wader chicks and suitable nesting sites | Positive |
| **H6** | Due to a lack of available food sources, peat bogs are not a key habitat for breeding waders | Negative |
| **H7** | Higher elevations are associated with lower food source availability per unit area, due to weather induced oligotrophic habitat such as peatland | Negative |
| **H8** | Coastal predators such as otters and gulls are likely to reduce breeding wader productivity, so distance from the the sea is positively associated with productivity. Coastal locations are possibly increasingly exposed to extreme weather events during the breeding season | Positive |
| **H9** | Bare peat in peatland habitats reduces the extent of suitable nesting and feeding habitats | Negative |
| **H10** | Available water capacity within the soil is an indicator of the degree of wetness of the soil which is positively associated with access to invertebrates in the topsoil | Positive |
| **H11** | Habitat with EUNIS classes D, E and F with high thematic complexity indicates high habitat heterogeneity, which is associated with higher diversity oand availability of nesting and feeding habitats in close proximity | Positive |
| **H12** | Habitat with EUNIS classes D, E and F with high spatially aggregation indicates high habitat homogeneity, which is associated with lower proximity of both nesting and chick feeding habitat | Negative |

Having quantified and evaluated the response of breeding wader abundance against the study hypotheses as outlined in Table 2, a population model will be developed to quantify the populations and spatial distribution of farmland breeding waders from 2002-2019, acros Shetland (excluding Foula and Fair Isle).

# Materials and Methods

## Shetland Breeding Bird Survey

The Shetland Breeding Bird Survey (BBS) is a citizen science project that is overseen by the Shetland Biological Records Centre (ref), the primary objective of which is to assess the population trends Shetlands more common breeding song birds and waders. The survey has been ongoing since 2002, where 36 local volunteers surveyed a total of 49 1km Ordnance Survey 1km squares. It is important to note that to encourage volunteer uptake, these squares are not randomly allocated in space; typically the location of the survey is selected by the volunteer. There are 3992 1km squares that cover the Shetland archipelago; excluding Fair Isle and Foula. A significant proportion of these squares only partially cover the landscape, hence why there are 3992 1km squares to cover the 1,466 km area of Shetland.

When under taking the Shetland BBS volunteers walk either a 2km transect around their chosen square, or two 1km transects that bisect the square. All breeding bird species observed 100m either side of the transect are recorded by the volunteer. All transects were walked twice, typically April to mid-May and mid-May to late June. All surveys were undertaken before 9am and in good weather (wind force four or less, and dry). The annual abundance recorded for each breeding species within a given square is the maximum count from the two separate visits. Upto 2019 there have been 139 different OS 1km squares surveyed as part of the SBBS. Although the survey runs every year, not all 139 survey squares have been covered annually. Since 2002, new squares have been introduced over the years, some are no longer covered, and there are some gaps when volunteers have been unable to carry out surveys.

## Estimating detectability

The Shetland BBS assumes perfect detectability of the survey species. Due to the imprecise nature of field surveys this could result in a bias in species detectability, which could further result in inaccurate population estimates. To account for this bias, there are a number of more sophisticated survey techniques such as *distance sampling* that account for this, in order to reduce the standard error associated with a population estimate (Newson et al. [2008](#ref-Newson2008-oq)). Such techniques typically require more structured survey techniques such as precisely measuring the distance and angle, between the target species and the transect line. Unfortunately this data was not available for the two survey datasets used within the study. In-order to generate a probability of detection, and therefore account for detectability bias, the r package unmarked (Fiske and Chandler [2011](#ref-Fiske2011-xf)) was used to generate a probability of detection for each species. Unmarked adopts a generative modelling process whereby observations are modelled through a combination of (1) a state process determining abundance at each site and (2) a detection process that yields observations conditional on the state process. Probability of detection (detectability) was modelled using unmarked for each species by using count data across multiple sites and years, using day of year as a covariate. An average was then taken across years to give an average detectability for each species.

## Exploratory Data Analysis

An exploratory data analysis (EDA) was undertaken on the Shetland BBS dataset as per the protocol outlined in Zuur (Zuur, Ieno, and Elphick [2010](#ref-Zuur2010-kp)). Applying basic EDA tests and visualisation techniques on to the survey count data helps to avoid rejection of a true null hypothesis, or type I error, or non-rejection of a false null hypothesis, a type II error. This analysis was also used to help formulate the survey hypotheses as set out in Section 1.6. The scope of the EDA undertaken comprised the following:

1. The survey effort (the number of surveys undertaken for a given 1km square) were plotted spatially in order to gauge survey effort, across the 139 different 1km squares that were surveyed between 2002 and 2019
2. Visualisation of the status of population change of each 1km square surveyed between 2002 and 2019 was undertaken, by species. This provides a simple overview of any significant population trends overtime.
3. A cleveland dot-plot was generated for the count data of each species, in-order to spot any outliers.
4. The count data were tested for normality, Poisson distribution fit and zero inflation. This enables the correct error distribution to be selected when selecting a suitable configuration for fitting regression model.
5. Homogeneity of variance was tested as this is an important assumption in certain regression model techniques.
6. As volunteers select the 1km square they want to survey, it is probable that the overall survey will be biased towards habitat that is easily accessible by volunteers. To assess the degree of any bias for grassland, heathland and bog habitats a bootstrap of habitat coverage was undertaken across the surveyed squares and compared to a bootstrap across all (n=3992) 1km squares across Shetland.

## Environmental covariates

Environmental covariates were generated with values spatially assigned to each Ordnance Survey (OS) 1km square (n=3992). The set of environmental covariates were also spatially joined to the Shetland BBS survey data. The spatial extent of the covariate grid excluded the remote islands of Foula and Fair Isle in order to reduce computational time when fitting spatial abundance models. The sub-sections below detail the environmental covariates that were generated for each of the 1km squares across Shetland. A spatial visualisation for each environmental covariate generated below can be seen in Appendix ??.

### EUNIS habitat classification raster

The EUNIS habitat classification is a European wide system for habitat classification (EUNIS [2019](#ref-Eunis2019-dy)). EUNIS classification habitat data for all of Scotland is available as a raster (Hijmans [2020](#ref-raster)); a data format used to store spatial data. For the purposes of EUNIS classification, a ‘habitat’ is defined as:

*‘a place where plants or animals normally live, characterized primarily by its physical features (topography, plant or animal physiognomy, soil characteristics, climate, water quality etc.) and secondarily by the species of plants and animals that live there’* (EUNIS [2019](#ref-Eunis2019-dy))

Each habitat type is identified by a specific code that is hierarchical and comprises three levels. For the purpose of this study the level 1 classification was used so as not to over-disperse the response variable (breeding wader abundance). The EUNIS dataset is available as a categorical raster at a resolution of 10x10m. The level 1 EUNIS categorisation is shown in Table 3.

Table 3: EUNIS Level 1 habiat classificaiton

|  |  |
| --- | --- |
| L1 Classification | Habitat Description |
| A | Marine habitats |
| B | Coastal habitats |
| C | Inland surface waters |
| D | Mires, bogs and fens |
| E | Grasslands and lands dominated by forbs, mosses or lichens |
| F | Heathland, scrub and tundra |
| G | Woodland, forest and other wooded land |
| H | Inland unvegetated or sparsely vegetated habitats |
| I | Regularly or recently cultivated agricultural, horticultural and domestic habitats |
| J | Constructed, industrial and other artificial habitats |
| K | Montane habitats |
| X | Habitat complexes |

A raster file containing the EUNIS habitat classification for Scotland in 2019 was cropped to the extent of the Shetland archipelago, and the percentage coverage of habitat classes *D, E and F* (bog, grassland and heathland) were calculated for each 1km square (n=3992). Wader breeding sites are assumed to be predominantly within D, E and F habitats.

### Improved grassland habitat classifiction

Across Shetland crofters have *improved* parcels of low-lying grassland in-order to take an annual crop of haylage or silage. It is typically used as fodder for sheep and cattle during the long winter season. This habitat is an important environmental covariate for breeding waders, as it has been suggested that lowland improved grasslands are an important feeding habitat for nesting waders and their chicks (McCallum et al. [2018](#ref-McCallum2018-gx)). Improved grassland is not a specific category within the EUNIS habitat classification, rather it is covered within the grasslands class (E), which in-turn also covers unimproved and rough grasslands. Consequently remotely sensed satellite image data were used together with supervised machine learning techniques (Abdi [2020](#ref-Abdi2020-aa)) to generate a spatial classification for improved grassland across Shetland.

The Sentinel-2 Level 2A product was downloaded for the period covering quarter three of 2019, and cropped to cover the Shetland archipelago. The raster dataset comprised a set of 11 spectral bands (ESA [2020](#ref-Esa2020-tl)) each with a spatial resolution of 10 m x 10m. The spectral bands layers used within the study for supervised classifiction of improved grassland are shown in Table 4.

Table 4: Sentinel 2 MSI Level 2A spectral bands used to train a support vector machine classifier for improved grassland habitat

|  |  |  |
| --- | --- | --- |
| Band ID | Name | Wavelength (micrometer) |
| 1 | Aerosol | 0.443 |
| 2 | Blue | 0.490 |
| 3 | Green | 0.560 |
| 4 | Red | 0.655 |
| 5 | Veg Red 1 | 0.705 |
| 6 | Veg Red 2 | 0.865 |
| 7 | Veg Red 3 | 0.740 |
| 8 | NIR | 0.783 |
| 8A | Veg Red 4 | 0.842 |
| 11 | SWIR 1 | 1.610 |
| 12 | SWIR 2 | 2.190 |

The r programming language (R Core Team [2019](#ref-R-lang)) was used to train a support vector machine (SVM) (Abdi [2020](#ref-Abdi2020-aa)) in-order to classify each 10m x 10m cell within the raster. The SVM was implemented using the tidymodels package (Kuhn and Wickham [2020](#ref-tidymodels)), according to the spectral band values within it. Training datasets comprising polygons were created across six different habitat classes: improved grassland, unimproved grassland, upland (heathland), bare peat, crops and cliff. Additional classes beyond improved grassland were created in order to disambiguate areas around improved grassland. A set of 100 raster cells at 10m x 10m were randomly sampled for each class from the training dataset; a sample comprised a value for each of the 11 spectral bands, for the given sample location. The aggregate training dataset, comprising 600 samples, was then split into training and testing data. The training data was subsequently resampled using 5-Fold cross-validation. Cross-validation seeks to evaluate model predictions by splitting data into a training sample and a validation sample; a candidate model is fitted with the training sample and evaluated against the validation sample. The hyper-parameters of the SVM were tuned using a regular tuning-grid, and the best model was selected based on the proportion of test data that were predicted correctly (model accuracy). The final model fit was then used to make a prediction for improved grassland over the entire Shetland archipelago raster.

### Median topsoil pH content

A shape file containing polygons of the median topsoil pH content across Shetland was downloaded by the James Hutton Institute (Donnelly, D and Buckler, C [2014](#ref-Donnelly_D_and_Buckler_C2014-hw)). A higher soil pH has been shown to be associated with greater earthworm abundance [McCallum2016-jt], which is one of the main food sources of breeding farmland waders and their chicks. The weighted median pH was calculated for each SBBS OS 1km square at a resolution of 10m x 10m, using the spatial features package (Pebesma [2018](#ref-spatial-features)) within r .

### Topsoil organic carbon content

A shape file containing the topsoil organic carbon content (by percentage of overall weight) was downloaded from the James Hutton Institue (Lilly, A, Baggaley, N and Donnelly, D [2012](#X5d4230c88f9295c6e868296d36d867050c13d20)). The topsoil of grasslands store significant amount of Carbon (Eze, Palmer, and Chapman [2018](#ref-Eze2018-pd)) through the accumulation of organic matter. Therefore fertile grasslands favoured by waders may have significant carbon stored within the topsoil, whilst grassland that is heavily grazed by livestock may have relatively low carbon storage.The mean topsoil organic carbon content was calculated for each Shetland BBS 1km square, using the spatial features package (Pebesma [2018](#ref-spatial-features)) within r.

### Available water capacity

Available water capacity (AWC) is the amount of water a soil can provide for plants, and so is a useful indicator of the ability of soils to grow crops as well as indicator of soil invertebrate density. For these reasons, the amount of moisture in the soil is thought to be an important aspect of wader abundance (Smart et al. [2006](#ref-Smart2006-dt)). A shape file detailing the AWC across Shetland was downloaded from the James Hutton Institute (Gagkas, Z., Lilly, A., Baggaley, N. & Donnelly, D. [2019](#Xc1035fe64dbdd93986e190df212fc806af046b3)), and the mean AWC across each Shetland BBS 1km square was calculated , using the spatial features package (Pebesma [2018](#ref-spatial-features)) within r.

### Bare peatland

Shetland has a significant amount of degraded peatland, due to natural wind and water erosion, worsened by past land management practices such as overgrazing and damaging methods of peat cutting for use as fuel. It is expected that the resulting bare peat is negatively associated with breeding waders. Scottish Natural Heritage have produced a shapefile of bare peat, at a resolution of 10x10m (Blake [2020](#ref-Blake2020-zj)). The percentage cover of bare peat for each Shetland BBS 1km square was calculated, using the spatial features package (Pebesma [2018](#ref-spatial-features)) within r.

### Distance to sea

Proximity of breeding wader territory to the coast as in-land breeding sites could be more sheltered from Shetland’s maritime weather. For each Sheltand BBS 1km square the mean distance form the center of the square to the coast was generated. This was generated using the spatial features package (Pebesma [2018](#ref-spatial-features)) within r.

### Elevation

The r package elevatr (Hollister and Tarak Shah [2017](#ref-elevatr)) was used to generate a mean elevation above sea level for each Shetland BBS 1km square. The purpose of this dataset was to establish if there was any statistically significant associations between breeding wader density in lowland habitat versus upland habitats.

## Bootstrap analysis of Shetland BBS survey squares

As Shetland BBS survey volunteers were able to choose which squares to survey, they are not randomly allocated. As such this may not be a true representation of the distribution of habitat types across Shetland. In order to analyse any potential bias in the habitat types surveyed, a bootstrap analysis (Davison and Hinkley [1997](#ref-bootstrap)) was performed across all Shetland 1km squares (n=3992), using the EUNIS habitat categorisation for grassland (E), heathland (F) and bog (D). This was then compared against a bootstrap of surveyed 1km squares Shetland BBS survey year (2018) in order to quantify how biased the volunteer survey samples are with respect to the overall Shetland habitat.

## Wader food response

## Breeding wader density response to environmental covariates

To investigate how breeding wader abundance responds to environmental covariates a generalised additive model (GAM) was used to fit wader abundance response (density per 1km square) to each environmental covariate. The GAM approach was chosen over linear regression modelling techniques because GAMs can effectively capture non-linear association, such as how covariates vary spatially. These effects can be captured using two dimensional non-parametric *smooths* (Wood [2003a](#ref-gams)) over longitude and latitude coordinates. Ecological count data are typically characterized by an excess of zeros and spatial dependence. The distribution of the response variable in a GAM can be parametrised as a zero-inflated poisson or negative binomial distributions, which are ideal for modelling species count data that are spatially distributed and with an excess of zeros.

The approach used in this section of the study is mostly based on the work undertaken to look at breeding Curlew density undertaken by the BTO (Franks et al. [2017](#ref-Franks2017-co)). Species count data from the SBBS was split into two time periods, each covering nine years; 2002-10 and 2011-19. Then for each species and each surveyed square within the two time periods, a mean count was derived for each of the two time periods and joined with the respective environmental covariates, as a function of spatial location. Breeding wader counts were modelled using a log link function with a *poisson* distribution in r using the gam function in the mgcv package (Wood [2003b](#ref-mgcv)). Location coordinates of survey squares were converted to planar OSGB36 national Grid coordinates as normalized easting and northing (in meters). Spatial variability in wader abundance was modelled as a non-parameteric spatial smooth, with the number of knots constrained (k = 20) and environmental covariates included as a parametric term. A different GAM was generated for each environmental covariate, in order to explore univariate associations between response and covariate variables. To explore absolute abundance of breeding wader species per 1km the natural log of the average detectability per species was included as an offset within each GAM. The GAM structure utilised is summarised in equation (1).

$$

, where *X* is the environmental covariate, *Y* is species count per surveyed square, *Z* is average species detectability and is random error. Associations between response and the spatial effect of environmental covariate were determined by the resulting p value of the gam fit. Only those models with p < 0.05 were considered statistically significant. Plots of density (abundance per 1km) against each environmental covariate were generated in order to visualise the trend of any association between response and environmental ovariate.

## Breeding wader population change as a response to environmental covariates

Given models for breeding wader density response against environmental covariates across two distinct analysis periods, a new set of GAMs were created in-order to model the change in abundance in 2002-10 relative to 2011-19. This was done by using the count in 2011-19 as the response, and the natural log of the 2002-10 count as an offset. It is assumed that detectability remained constant within each square, which in turn means that changes in count reflect changes in density, against a given environmental covariate. Only Shetland BBS squares that were surveyed in both periods and where breeding waders were present in 2002-10 where used, as it is not possible to take the natural log of zero. The population change ratio was then plotted against each environmental covariate, in order to determine the association of any population change as a function of each environmental covariate.

## Habitat spatial complexity response

In order to investigate an association between breeding wader abundance and the complexity of habitat spatial patterns within the landscape, metrics using information theory (IT) (Nowosad and Stepinski [2019](#ref-Nowosad2019-bo)). IT metrics enabled a way to numerically quantify habitat pattern complexity, for all Shetland 1km squares. Results from each metric were then incorporated into a GAM environmental covariate and population change models, in order to explore if varying habitat spatial patterns have any significant statistical associations with breeding wader abundance.

The r package landscapemetrics (Hesselbarth et al. [2019](#ref-landscapemetrics)) was used to generate all IT metrics, using the EUNIS habitat raster (EUNIS [2019](#ref-Eunis2019-dy)) for Shetland as a spatially distributed categorisation of landscape habitat. A description of the different metrics incorporated into response the analysis is as follows:

* *Marginal Entropy, H(X)* - this first metric measures the compositional diversity of habitats in space; from mono-thematic patterns (lower H(X)) to multi-thematic patterns (large (H(X)). So landscapes with one dominating habitat (for example, peatland on the Shetland island of Yell) have a relatively low H(X), whilst landscapes with multiple different types of habitat that are evenly distributed (such as a village with a mosaic of gardens, housing and undeveloped land) have a relatively high H(X). This can be seen in Figure 1, with H(X) shown above each landscape sample.

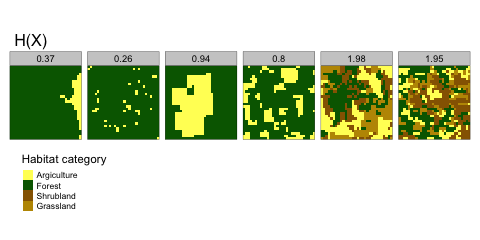


Figure 1: Marginal Entropy

* *Conditional Entropy, H(Y|X)* - quantifies the geometric intricacy of a spatial pattern within a landscape. If habitat type A is predominantly adjacent to habitat type B, H(Y|X) will be relatively low. Conversely if habitat type A is adjacent to many different habitat categories, then H(Y|X) will be relatively high. Figure 2 gives some examples of this.

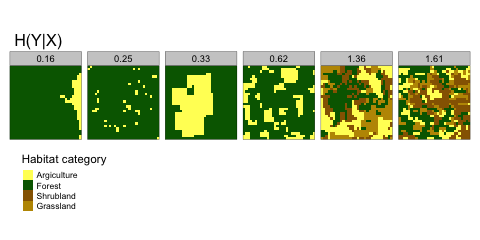


Figure 2: Conditional Entropy

* *Joint Entropy, H(X,Y)* - this provides a measure of the uncertainty in determining the habitat category of a focus cell and an adjacent cell. So landscapes with high H(X,Y) are typically spatially complex with many different habitat types. Note that joint entropy is not capable of distinguishing between patterns that have high spatial aggregation. The variation can be seen in Figure 3 .

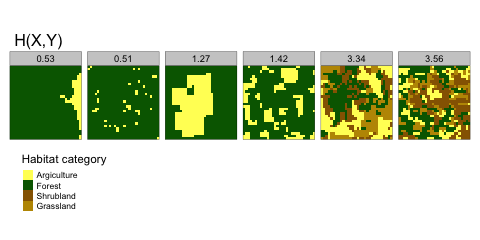


Figure 3: Joint Entropy

* *Relative mutual information, U* - quantifies the degree of aggregation (clumpliness) of spatial habitat categories from fragmented patterns (lower U) to consolidated patterns (higher U). A landscape comprising a loch within a forest would have a relatively high U, whilst a landscape comprising many different crop types spread across many small fields would have low U. Figure 4 gives a mutual entropy landscape

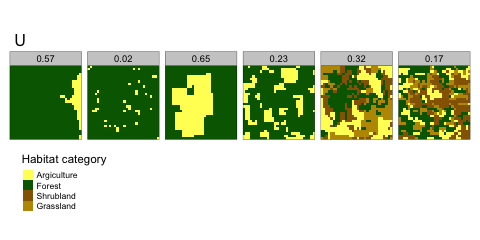


Figure 4: Relative Mututal Entropy

## Wader population abundance modelling

Given the set of environmental covariates for each Shetland BBS 1km square and the associated count data for those squares that were surveyed, a machine learning approach using the tidymodels (Kuhn and Wickham [2020](#ref-tidymodels)) package was taken to fit a *random forest* regression model (Andy Liaw and Matthew Wiener [2002](#ref-Liaw_A_and_WinerM2002-te)).

In order to fit the model the count data for each species were adjusted according to their mean detectability. All count data for a given Shetland BBS 1km square was joined with the associated environmental covariate data. The data were then pre-processed so that covariates that have 80% absolute correlations with other covariates are removed; this ensures that possible adverse outcomes due to co-linearity are minimised (Zuur, Ieno, and Elphick [2010](#ref-Zuur2010-kp)). The covariate data within the training dataset were then normalised to a mean of zero and a standard deviation of one. The joined data were then split such that 60% of all data were allocated to model training and the remaining 30% to model testing. A random forest machine learning package (Wright and Ziegler [2017](#ref-ranger)) was used to fit a regression model. 10-fold cross validation was used to resample the data in order to ensure any bias in Shetland BBS survey squares was minimised. Hyper-parameters used to fit the random forest model were selected using a grid containing 10 random variations for each hyperparameter:

* min\_n - the minimum number of data points in a node that are required for the node to be split further
* trees - the number of trees contained in the ensemble
* mtry - the number of predictors (covariates) that will be randomly sampled at each split when creating the tree models.

A model was fitted for each combination of hyper-parameters, for each species; giving 50 different models in total. The best model fit for each each species was selected according to the lowest root mean squared error (rmse). The hyper-parameters associated with each best fit were then used to further tune the model for a final fit. The rmse for the final fit was then evaluated against the test dataset, for each species. This gave an abundance estimate (per BBS km$^2^$) for each breeding wader species, for each survey year. Summing the density estimate across all squares, by species per year gives an annual mean population estimate for each breeding wader species across all of Shetland, from the 2002 to 2019. Using bootstrap aggregation or *bagging*, it was possible to generate a lower and upper confidence interval for each of the annual population estimates.

## Lapwing population association with reseeded grassland

Given the population abundance estimate method as outlined in Table 2.10, a specific association between the Lapwing population trend over time and agricultural census data (Scottish Government et al. [2003](#ref-Scottish_Government2003-gf)) was investigated. Specifically the size of different types of annual grassland holdings in Shetland were used as covariate in a regression analysis between 2002 and 2017. The data were tested for normality and it was found that they were log-normally distributed. A generalised additive model for location scale and shape was used to fit the response variable, Lapwing population estimates, against the size of various grassland holdings.

## Wader spatial desnity modelling

Having generated an density estiamtion for each Shetland BBS square, the spatial density can be plotted using the spatial coordinates for each SBBS square, and for each year. Spatial density were generated for each species for 2019 in order to visualise spatial abundance distribution. The net change between the two periods was also plotted so as to visualise the areas of Shetland that have seen a net change in species density between 2002 and 2019.

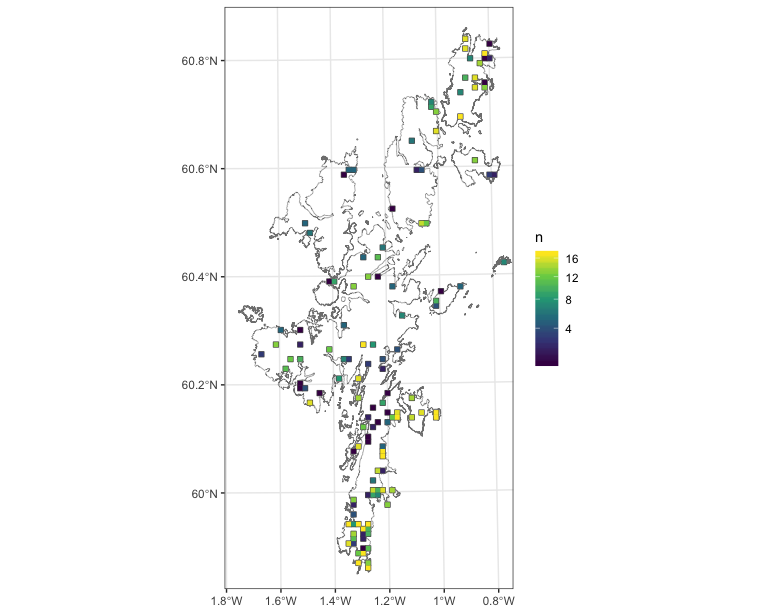
# Results

## Explororatory Data Analysis of Shetland BBS survey data

Where relevant a protocol for exploratory data analysis was followed (Zuur, Ieno, and Elphick [2010](#ref-Zuur2010-kp)) to ensure that before any problems in the structure of the data are identified prior to undertaking any statistical analysis.

### Survey effort over time

The spatial location of surveyed squares are shown in Figure 5. It seems that there has been ongoing surveying effort in the south mainland and on the islands of Unst, Bressay and Noss, but less coverage elsewhere. In particular bog and upland heathland have significantly less survey effort. This finding is also seen in the bootstrap of surveyed habitat versus overall Shetland habitat (see Section 3.1.6).

 ### Status of population change at survey sites between 2002 and 2019

Before any detailed statistical modelling was undertaken a simple analysis into how the population changed between 2002-2011 and 2012-2019, in each surveyed 1km square (n=139). This gives an initial view as to potential population trends between the two analysis periods.The 1 km squares included were those surveyed in both periods and where farmland waders colonized, increased, remained stable, declined or went extinct. Figure 6 shows the status changes between the two analysis periods.

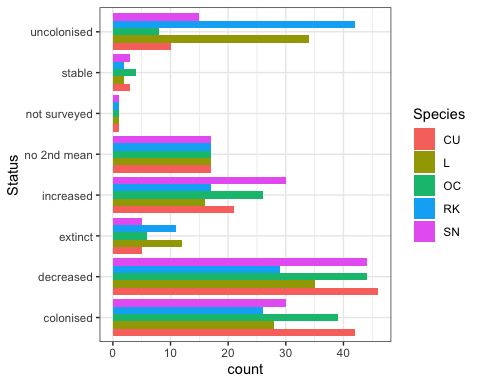


Figure 6: Population status change per Shetland BBS square - between 2002-10 and 2011-19

Figure 7 below shows an aggregation of certain categories whereby extinct and decreased are grouped, and colonised and increased are grouped. It can be seen that only Lapwing show a net decline over the two analysis periods.

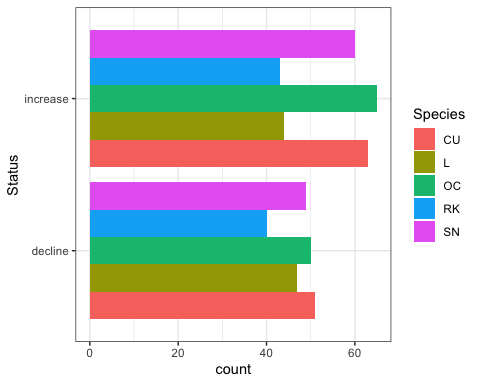


Figure 7: Aggregate population status change per Shetland BBS square - between 2002-10 and 2011-19

### Outliers

The *Cleveland dot plot* (Zuur, Ieno, and Elphick [2010](#ref-Zuur2010-kp)) is a chart in which the row number of an observation is plotted versus the observation variable, thereby providing a more detailed view of individual observations than a boxplot. Points that stick out on the right-hand side, or on the left-hand side, are observed values that are considerably larger, or smaller, than the majority of the observations. Figure 8 appears to show that there are no significant outliers across all species, but that there are many counts equal to zero indicating that the count data might be zero-inflated.

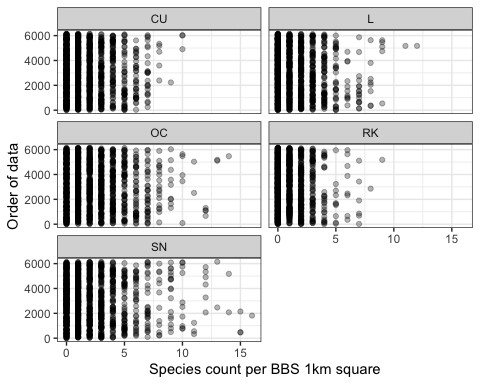


Figure 8: Cleveland dot plot of species counts in Shetland BBS data from 2002 to 2019

### Testing for normality

A large number of statistical regression techniques assume normality. Visualising the Shetland BBS count data as a histogram can help assess if it is normally distributed. This is shown in the plot in Figure 9.

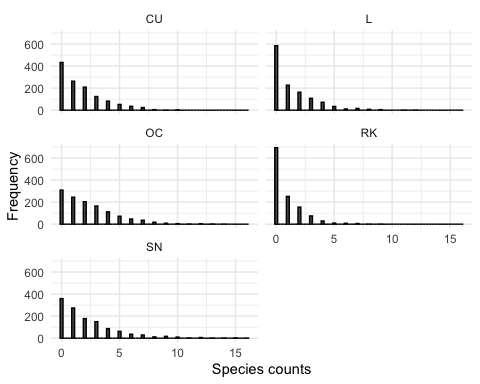


Figure 9: Histogram of SBBS count data across all years, by species

Clearly the count data are not normally distributed. In-order to validate the outcome of the plots in Figure 9 a Shapiro-Wilks normality significance test was undertaken and the results are shown in Table 5.

Table 5: Shapiro-Wilks normality test for count data of breeding waders

|  |  |  |
| --- | --- | --- |
| Species | W | p-value |
| OC | 0.8628058 | 0 |
| L | 0.7546102 | 0 |
| CU | 0.8327195 | 0 |
| RK | 0.6995970 | 0 |
| SN | 0.8000661 | 0 |

The p-value for each species in 5 is << 0.05. This suggests that the count data for all species are significantly different from the normal distribution.

### Poisson distribution and zero inflation

The histograms of species counts in Figure 9 suggest that count data is a poisson distribution. Also there are a significant number of zeros in the count data, across all species count data. This suggests that the zero-inflation poisson distribution describes the data. Table 6 below shows the results of a significance test (Broek [1995](#ref-Van_den_Broek1995-ml)) for data zero inflation that follow a poisson distribution.

Table 6: Zero inflation poisson distribution test for Shetland SBBS count data, across all years and species

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Species | Expected zeros | Zeros observed | Chi squared | p-value |
| CU | 223.9173 | 433 | 384.2445 | 0 |
| L | 323.5230 | 585 | 547.5071 | 0 |
| OC | 119.1142 | 309 | 446.9641 | 0 |
| RK | 520.3458 | 694 | 271.6801 | 0 |
| SN | 134.4406 | 359 | 577.9990 | 0 |

All results have a significant statistical significance (p<0.05) and therefore the count distribution across species data are assumed to be a zero-inflated poisson process. The statistical modelling methods used on the data must support a poisson distribution and zero inflation where possible.

### Homogeniety of variance

Homogeneity of variance within the data is an important assumption in analysis of variance (ANOVA) and other regression-related models. The series of boxplots in Figure 10 show how counts across all surveyed BBS squares vary across years 2002 to 2019, for each breeding wader species.

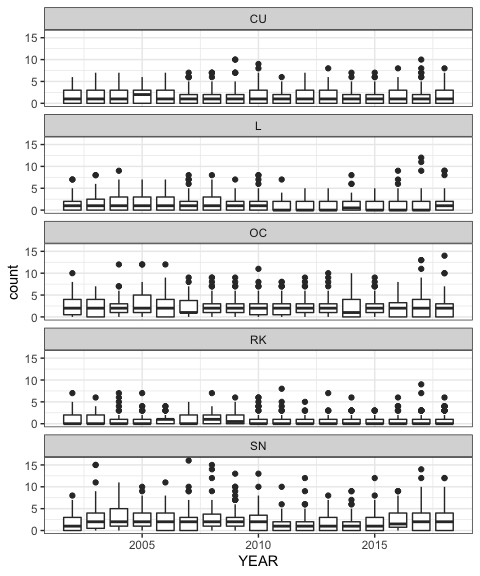


Figure 10: Box plot showing variance of counts across all surveyed Shetland BBS squares and all years, by species

To test the homogeneity of variance of species counts between years, for each species, we can apply the Fligner-Killeen test. This is used as the count data are shown to be non-normal. Table 7 shows the results of the test applied to the Shetland BBS data. For p-values > 0.05 the data variance are homogeneous.

Table 7: Fligner-Killeen test of homogeneity of variance for Shetland SBBS species counts, across all years

|  |  |  |  |
| --- | --- | --- | --- |
| Species | Chi-squared | p-value | df |
| OC | 18.11614 | 0.3171401 | 16 |
| L | 29.11514 | 0.0231712 | 16 |
| CU | 18.36325 | 0.3030590 | 16 |
| RK | 26.72879 | 0.0445979 | 16 |
| SN | 47.40824 | 0.0000588 | 16 |

Lapwing, Redshank and Snipe variances are heterogeneous according to the test results in Table 7. The solution to heterogeneity of variance is to transform the response variable to stabilize the variance year-on-year, or applying statistical regression techniques that do not require homogeneity, such as a generalised additive model.

### Survey Bootstrap

Shetland BBS volunteers were able to choose which squares they surveyed. As a result of this non-randomised allocation there could be potential bias in the habitat types surveyed; for example, in-by is closer to roads and housing than upland habitats. To test this a bootstrap of percentage cover of EUNIS habitat categories D, E and F (see 3) across all Shetland 1km squares was undertaken, and then compared to a bootstrap of the same data, but only those squares surveyed by volunteers as part of the Shetland BBS.

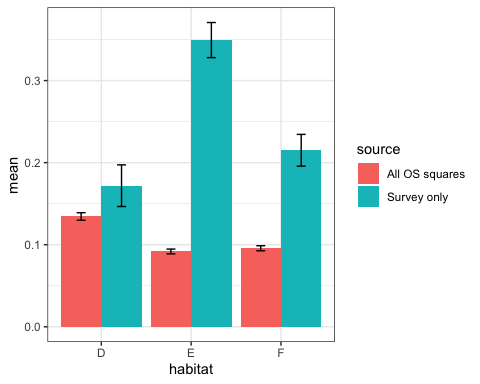


Figure 11: Mean % cover per 1km of EUNIS habitat types D, E and F, bootstrap sample of OSGB squares v boostrap of surveyed squares. R=1000

This shows that grassland and heathland are significantly oversampled within the Shetland BBS surveys, but bog habitats appear to be sampled proportionally.

## Detectability

The r package unmarked was used to generate an estimate for the probability of detection, or *detectability*. Table 8 shows the average *detectability* across all survey years, for each species.

Table 8: Average detectability of breeding wader species on Shetland 2002-2019

|  |  |
| --- | --- |
| Species | Detectability |
| OC | 0.803 |
| L | 0.723 |
| RK | 0.667 |
| CU | 0.831 |
| SN | 0.723 |

## Improved grassland classification

The results below detail how remotely sensed satellite data was processed using the Support Vector Machine (SVM) machine learning algorithm to generate a classification for improved grassland across Shetland.

### Shetland Sentinel 2 satelitte dataset

A Sentinel 2 Level 2A spatial dataset was clipped using the Integrated Administration And Control System (IACS) (Oesterle and Wildmann [2003](#ref-Oesterle2003-zb)) field boundary shapefile for Shetland. This gave a spatial dataset comprising of land-based habitat only, excluding built areas and roads. The RGB composite of the clipped image is shown in Figure 12

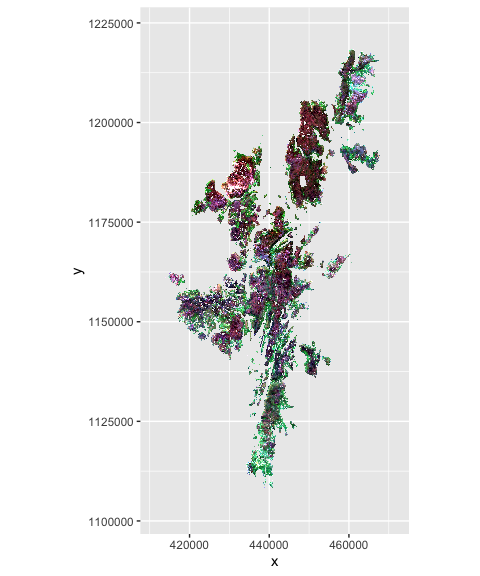


Figure 12: Clipped Sentinel 2 RGB composite of Shetland

### Habitiat classification training data

In order to classify improved grassland, five other distinctive habitat types were also classified: unimproved grassland, crops, bare peat, cliffs and upland. A number of areas representative of each habitat type were selected as shown in Figure 13.

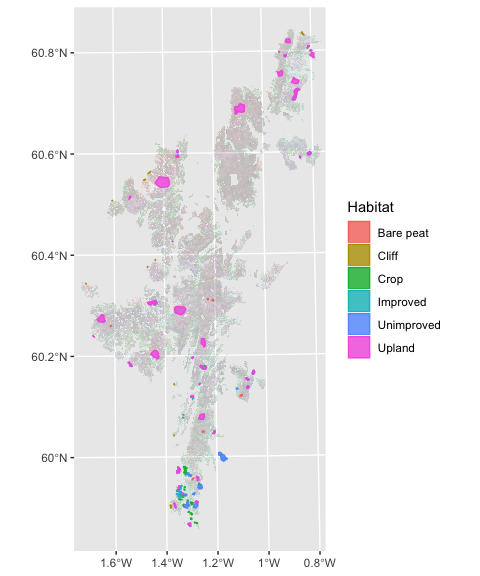


Figure 13: Habitat classification training areas

### Sampling of habitat training classes

Each habitat training dataset was randomly sampled in order to train the SVM habitat classifier. Distributions for the sampled data for each training set are shown in Figure 14.

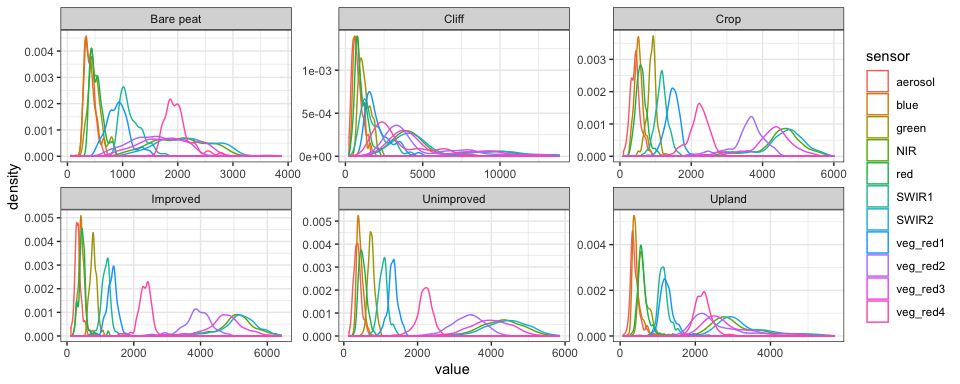


Figure 14: Sampled distributions for each training class, from the Sentinel-2 Level 2A dataset of the Shetland archipelago

There are a number of distinguishing observations that can be made from Figure 14. Firstly a visual inspection of the bare peat, cliff, upland and crop classes show a clearly different spectral finger-print for each habitat type. Perhaps unsurprisingly improved and unimproved grassland are relatively similar. On closer inspection it can be seen that the shape of the near infra-red (NIR) spectral histogram appears significantly different for improved grassland. A significant proportion of NIR light is reflected by green vegetation, therefore greener “improved” vegetation is likely to have a strong NIR spectral response (Pettorelli et al. [2014](#ref-Pettorelli2014-ad)).

### Support vector machine classifier training

In order to undertake supervised training of an SVM, optimal values for for so-called *hyper-parameters* must be selected. The type of SVM tuned for classifying improved grassland was a radial basis SVM, and two key hyper-parameters must be tuned for when selecting the best model fit. A so-called *search grid* is used to specify all permutations of hyper-parameters that will be utilised to find the optimum fit. For the improved grassland classification, the specific search parameters used are shown in Table 9.

Table 9: Search parameters used for model fitting of support vector machine

|  |  |
| --- | --- |
| RBF sigma | Cost |
| 0.11 | 18 |
| 0.12 | 18 |
| 0.13 | 18 |
| 0.11 | 19 |
| 0.12 | 19 |
| 0.13 | 19 |
| 0.11 | 20 |
| 0.12 | 20 |
| 0.13 | 20 |

### Best model results by classification accuracy

The best SVM model hyper-parameters, as measured by classification accuracy, are shown in Table 10. It can be seen that when the best model is used to make classification predictions using the training, it achieved an accuracy of 82% with a standard error of 1%.

Table 10: SVM model parameters from best fitting models

|  |  |  |  |
| --- | --- | --- | --- |
| Cost | RBF sigma | Accuracy | se |
| 20 | 0.13 | 0.8233010 | 0.0094407 |
| 19 | 0.13 | 0.8208738 | 0.0093362 |
| 18 | 0.13 | 0.8194175 | 0.0093179 |
| 19 | 0.12 | 0.8194175 | 0.0091478 |
| 20 | 0.12 | 0.8194175 | 0.0090903 |

### Model performance using the test dataset

Having selected the best model fit, the same model was used to make predictions against the test dataset. The SVM classifier accuracy was shown to be 85% as shown in Table 11.

Table 11: Classifier performance using test data set

|  |  |
| --- | --- |
| Metric | Estimate |
| accuracy | 0.8534091 |
| kap | 0.8240900 |

The confusion matrix in Figure 15 shows the results of the model prediction for each habitat class against those of the test data set. The most incorrectly classified habitat is unimproved grassland (class 2) versus upland (class 3), followed by improved grassland (class 1) versus unimproved grassland. This gives an accuracy for improved grassland classification of c.85% over the test dataset.

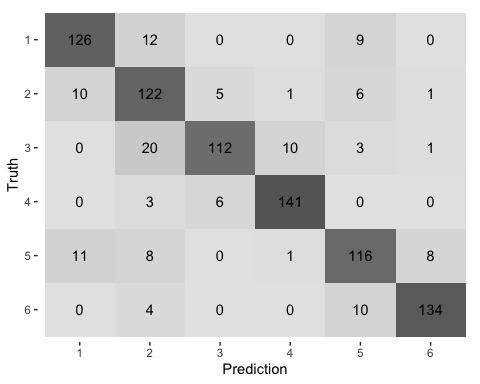


Figure 15: Confusion Matrix from classificaton of test dataset

### Classification across all Shetland habitat

The best fitting model was used across the entire raster dataset for Shetland, to enable classification of all habitat according to the chosen six habitat classes. The results are shown in 16. It can be seen that the improved grassland and crop habitat classes are predominantly in the south of the islands; this was validated by a local expert. It can be seen that the main middle island, Yell, is predominantly upland and bare peat.

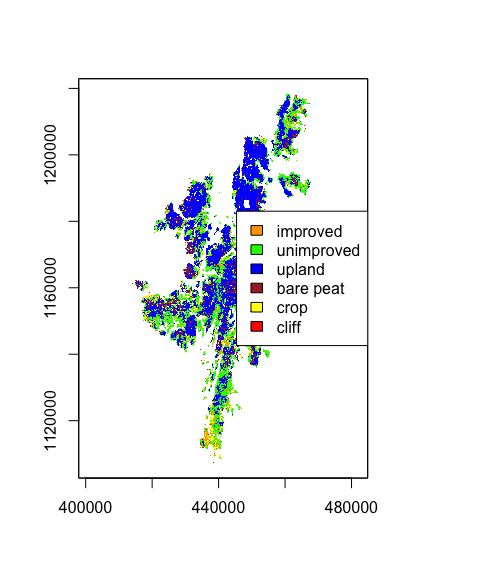


Figure 16: Classification of Shetland habitat in order to determine the location of improved grassland

## Environmental covariate analysis

Each of the covariates described in Section 2.4 were generated for each OS 1km Shetland square (n=3992).

### Histogram of environmental covariates

Figure 17 shows histograms for all environmental covariates across the Shetland archipelago, at 1km resolution. It can be seen that bog type habitat predominates across Shetland, and that majority of the landscape is at less than 100m elevation. The mode for the pH is around 4, which is typical for acidic peatland (Paterson [2011](#ref-Paterson2011-ky)). The topsoil carbon content shows that the majority of the Shetland soils have a high organic carbon content; again this is typical of peatland (Paterson [2011](#ref-Paterson2011-ky)). In contrast a typical mineral rich soil in southern England would have c.3-5% organic carbon content. Note that the percentage landcover of improved grassland is relatively low compared to the more general grassland classification. This shows that agricultural intensification is relatively low in Shetland.

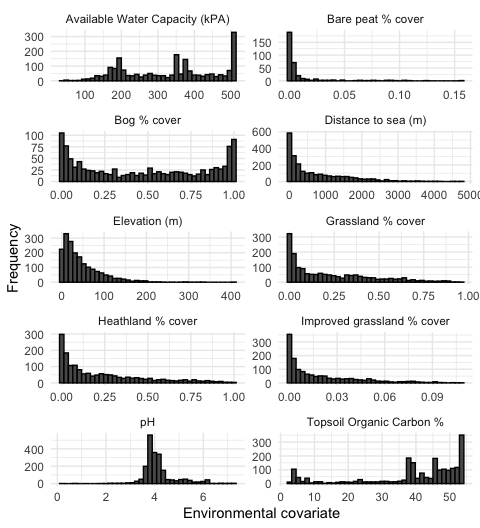


Figure 17: Histograms of environmental covariates across all of Shetland

### Histogram of environmental covariates for Shetland BBS squares only

This can be contrasted with covariate histograms for only those OS squares (n=139) that were surveyed as part of the Shetland BBS, as seen in Figure 18. If these data are indicative of general breeding wader habitat preferences, it would seem that across all surveyed nesting wader species, there is a preference for wet but not water-logged habitat as seen in the AWC histogram. Also, the majority of breeding waders that were surveyed appear to nest within 1km of the coast. It appears that surveyed breeding waders also have a preference for grassland (both improved and unimproved) presents the majority of the habitat cover, over heathland.

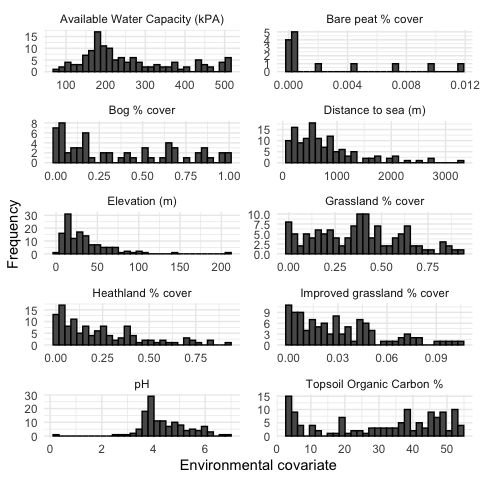


Figure 18: Histograms of environmental covariates across only those squares surveyed as part of the Shetland BBS

### Density plots of environmental covariates

Density plots of all environmental covariates across Shetland BBS squares (n=139) are shown in Figure 19. These figures provide an overlay to the histograms in Section 3.4.2, and represent a smoothed version of a histogram to show the probability density function of the variable. Some distributions are highly skewed, such as distance to sea and elevation. Whilst other covariates like topsoil organic carbon and bog cover are largely a uniform distributed. None of the covariates appear to be normally distributed.

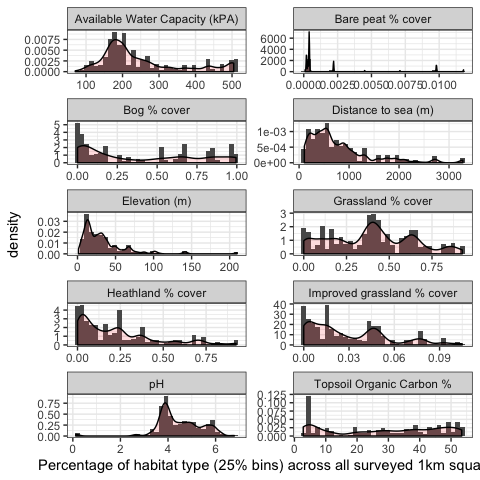
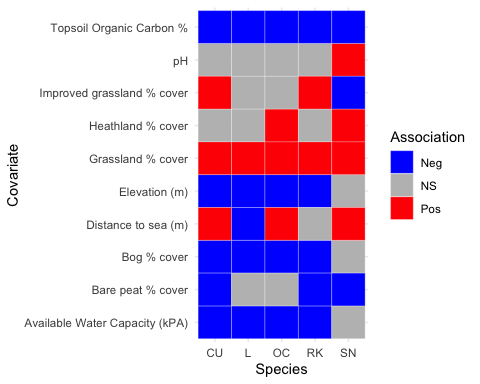


Figure 19: Density plots of environmental covariates against breeding wader count data

## Breeding wader abundance response to environmental covariates

Generalised Additive Models (GAMs) for breeding wader abundance response across the five different species were generated from 2002–10 and 2011-19. A GAM was produced for each of the 10 environmental covariates (predictor variables), and for each of the two analysis periods.

Appendix ?? details the GAM parameter results from the model fitting process, for the two periods where the abundance response (density) was modelled against environmental covariates. The associated plots showing breeding wader density response against each environmental covariate are shown in Appendix ??. Note those models that result in a parametric term (an environmental covariate) that is statistically significant (p<0.05) have plots that are coloured red. The statistically significant correlations between breeding wader density and environmental covariates are summarised for each species in heatmaps; Figure 20 for 2002-10 and Figure 21 for 2011-2019.

 For the 2002-10 survey period in Figure 20, it can be seen that pH is only statistically significant for one species (Snipe), whilst topsoil organic carbon and grassland are oppositely correlated. Distance to sea is perhaps the most interesting covariate in that Lapwing show a greater association at the coast whilst Curlew, Oystercatcher and Snipe show greater densities inland. Notable results are that all nesting species are negatively associated with higher percentages of carbon within the topsoil per km, and all nesting species are positively associated with increased grassland coverage per km.

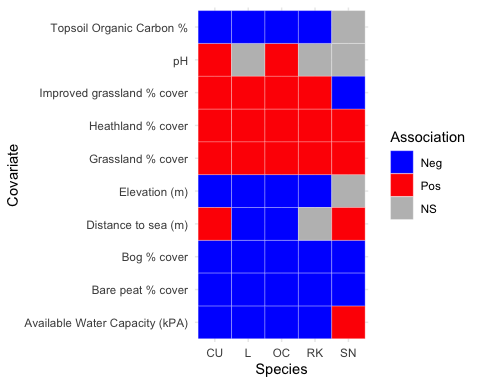


Figure 21: Summary of associations between breeding wader density and environmental covariate, between 2011 and 2019 inclusive

In Figure 21 we can see that there are fewer associations that are not statistically significant. Again all species are positively associated with increased grassland coverage per km, and in the second analysis period this is true for increased heathland coverage per km. All species are now negatively associated with increased bog coverage per km and increased bare peat coverage per km. It can also be seen that all species, apart from Snipe, are positively associated with increased improved grassland coverage per km. Where as the exact opposite is true for available water capacity per km.

For Curlew it is seen that all covariate associations are now statistically significant (versus 2002-2010 where pH and Heathland cover were not significant). Oystercatcher also have all covariates with statistically significant associations. Note that for Oystercatcher, for the second analysis period, results for the distance to sea association show a negative association with increasing distance. Lapwing only have one covariate, pH, that is not statistically significant, and otherwise show identical results to Oystercatcher. For Redshank the main change in the later analysis period is that Heathland percentage cover is now statistically significant, and has a positive association. Snipe now have a positive association with available water capacity and a negative association with percentage bog cover.

## Population change response against environmental covariates

A third model was generated by using the abundance response from the first two GAMs in Section 3.5. By using the response of the 2002-2010 wader densities as the offset for the 2011-19 densities, a third series of GAMs were fitted to show the ratio of population change as a function of environmental covariates. The population change model will give an indication of how breeding wader densities have changed over time for a given environmental covariate. For example, are species decreasing over time in habitats that are thought to offer lower food availability for chicks, such as heathland.. The results of the analysis are shown in Figure 22.

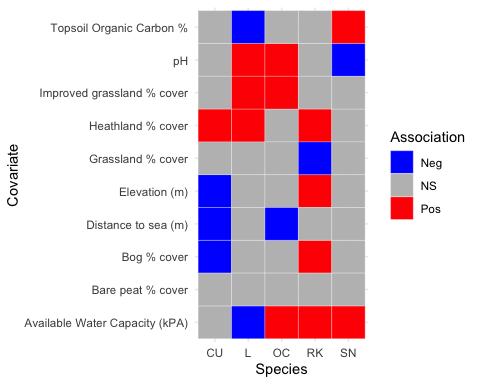


Figure 22: Summary of population change ratio associations between breeding wader density and environmental covariate, between 2002 and 2019 inclusive

Figure 22 suggests that the environmental covariates that have had the most positive associations with breeding wader population changes over the two analysis periods are heathland percentage cover and available water capacity. The percentage of bare peatland has had no statistical significance, followed by the percentage grassland cover that has only one negative association with the population change ratio of Redshank. The model parameters and associated plots for population change ratio modelling are shown in Appendix ?? and Appendix ?? respectively.

## Information Theory (IT) covariates

GAMs for breeding wader abundance response across the five different species were generated from 2002–10 and 2011-19. In addition to the environmental covariates, they included five IT covariates. The model parameter results for all species across the two analysis periods can be seen in Appendix ??. A third model was generated by taking the abundance response from the first two models to generate the ratio of population change between the two response periods, the model parameters for this model can be seen in Appendix @ref{#gam-it-pop-chg-params}.

### Histograms of Information theory covariates

Histograms of IT covariates using the EUNIS landscape categorisation, across all of Shetland are shown in figure 23. The *marginal entropy* for the Shetland landscape is approximately normally distributed, indicating that habitat within the Shetland landscape is spatially diverse but that very low and highly diverse habitat within Shetland are relatively rare. The mode of the *conditional entropy* is relatively low with a distribution that shows significant positive skew; this suggests that the Shetland landscape has relatively low geometric intricacy. This arises when cells of one category within a landscape raster are predominantly adjacent to cells of the same category. The overall spatio-thematic complexity is measured by the *joint entropy*. This can be thought of as quantifying the uncertainty in determining the habitat type of a focus cell and an adjacent cell. For Shetland, joint entropy appears to be approximately normally distributed. This indicates that habitat with very high or low spatio-thematic complexity is relatively rare on Shetland. Due to the spatial autocorrelation, the value of *mutual information* tends to grow with a diversity of the landscape (marginal entropy). To adjust this tendency, it is possible to calculate *relative mutual information* by dividing the mutual information by the marginal entropy. Relative mutual information always has a range between 0 and 1, and quantifies the degree of aggregation of spatial habitat. It can be seen that for Shetland, relative mutual information is distributed with significant negative skew. This implies that habitat types across Shetland are predominantly aggregated - small relatively information values indicate significant fragmentation in landscape habitat patterns.

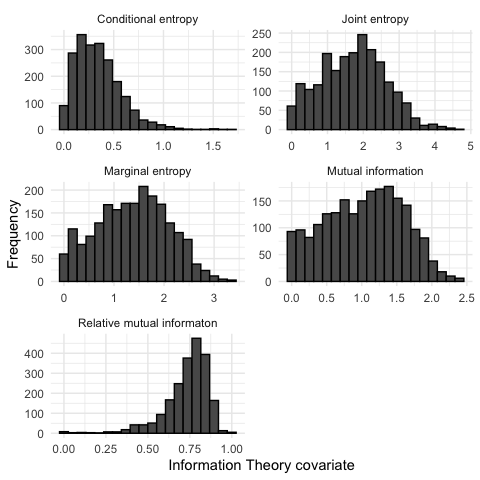


Figure 23: Histograms of Information Theory covariates across all Shetland OS 1km squares

### Histograms of Information Theory covariates for surveyed squares only

1km squares surveyed as part of the Shetland BBS were used to generate IT covariates histograms using the EUNIS landscape categorisation, as shown in Figure 24. Here we can see that the *conditional entropy* and the *marginal entropy* across all surveyed squares had a mode that was significantly higher than the Shetland wide values shown in Figure 23. There is also significantly less negative negative skew in the *relative mutual information* of the surveyed squares.

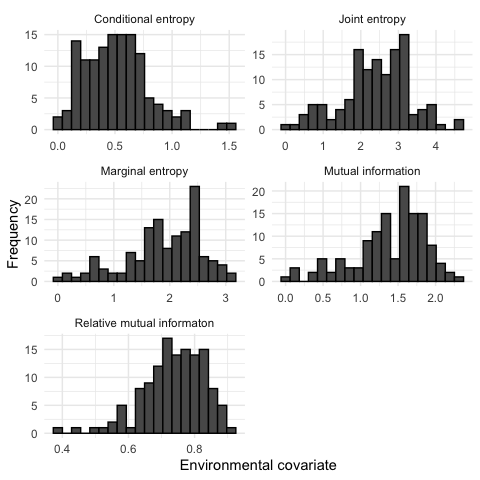


Figure 24: Histograms of Information Theory covariates across SBBS surveyed sqaures only

### Information Theory covariates abundance response model

GAMs were fitted for each of the two analysis periods using IT metrics as covariates against breeding wader abundance. Figures 25 and 26 summarise the associations between the abundance response and IT covariates used in the univariate GAMs, for the periods 2002-10 and 2011-19 respectively.

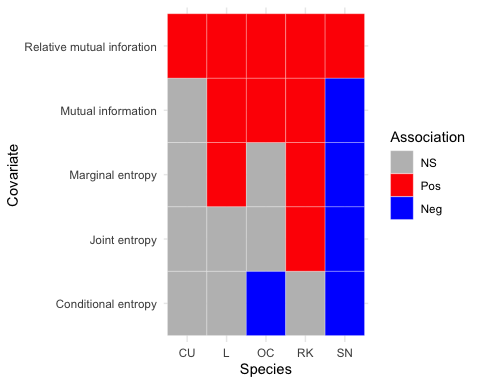
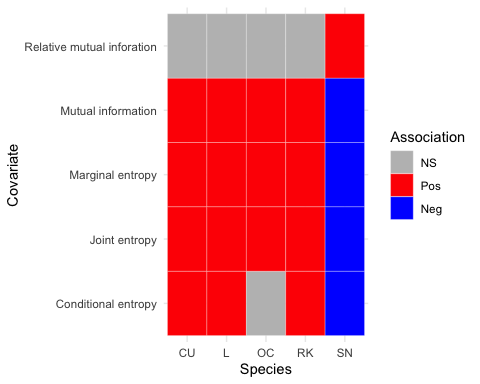


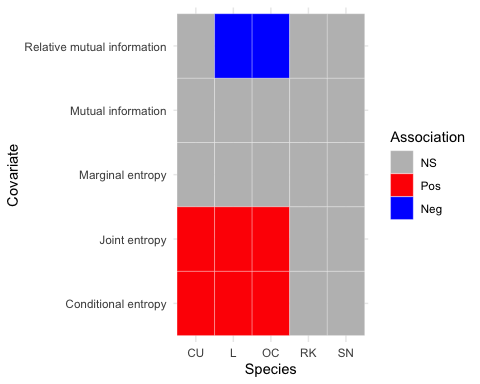
Figure 25: Summary of associations between breeding wader density and IT covariates, between 2002 and 2010 inclusive

 Appendix ?? shows the GAM parameter results generated by fitting the model to the data when information theory metrics are used as covariates. Plots showing the abundance response against IT covariates are shown in Appendix ??.

It seems that the response for Snipe abundance to IT covariates is identical for the two periods. Snipe have a positive association with *relative mutual information*, or habitats that have a high degree of aggregation, such as heathland. *Relative mutual information* has no statistically significant association for any other species in the second analysis period, although it appears to be statistically significant and positively associated for all species in the first analysis period. This indicates that breeding waders, apart from Snipe, may have moved from areas comprising habitat that is highly aggregated to areas that are less so. This hypothesis is partially supported by the fact that the second analysis period shows increase in wader species, apart from Snipe, that are positively associated with *marginal entropy*; a measure of habitat thematic diversity.

### Population change model against IT covarirates

By using the response of the 2002-2010 wader densities as the offset for the 2011-19 densities, a third series of GAMs were fitted to show the ratio of population change in response to IT covariates. This is summarised in Figure 27. It can be seen that there are no statistically significant results for Redshank or Snipe, or for marginal entropy as a covariate.

 Perhaps the most significant result is that which supports the idea that breeding wader densities have declined in habitats that exhibit a large spatial aggregation. This is shown by the fact that Lapwing and Oystercatcher populations are negativelty associated over time, with increased *relative mutual information*. Appendix ?? shows the GAM parameters generated by fitting the population change model to the data. Plots for population change response against IT covariates are shown in Figure ??.

## Wader abundance trends

A random forest regression model was used to fit an abundance response using all 10 environmental covariates. The parameterisation and results are outlined below.

### Tuning the model hyper parameters

A search grid over 10 different folds gave initial results for various permutations of hyper parameters as shown in Appendix ??. Each figure shows the results for a particular species. Each hyper-parameter permutation is plotted against the resulting root mean squared error (rmse).

### Further model hyper parameter tuning

Given the initial results in ?? the initial range over which the hyper-parameter search was conducted, was refined further by searching over a revised range according by selecting upper and lower limits for hyper-parameters that gave the lowest rmse as indicated by the initial results. The range used was that which gave the lowest rmse as given in Appendix ??. The results of the revised search grid range can be seen in Appendix ??. Again each plot is for a separate species.

From the plots in Appensix ?? we can see which hyper parameters give the best fit, when using root mean squared error as an evaluation metric. It can be seen that the model fit for Snipe has the largest RMSE. whilst Redshank has the lowest. For each species the minimum rmse result for the best model fit, together with the associated hyper parameters (trees= 1000 for all models) is shown in Table 12.

Table 12: Lowest RMSE for best fitting model, for each species of breeding wader

|  |  |  |  |
| --- | --- | --- | --- |
| Species | mtry | min\_n | rmse |
| CU | 4 | 6 | 1.448835 |
| L | 5 | 4 | 1.672523 |
| OC | 4 | 10 | 1.845994 |
| RK | 7 | 3 | 1.429948 |
| SN | 2 | 6 | 2.718174 |

### Variable importance in model fit

Having selected the best model fit the variable importance for each species was assessed. The r package vip (Greenwell, Boehmke, and Gray [2020](#ref-vip)) was used to explore the relative importance of different covariates in the model fit. The results are shown in Appendix ??.

It can be seen that pH, X (longitude) and grassland percentage coverage for a given OS 1km square appear to be the most important covariates for predicting abundance in Curlew. For Lapwing, pH, heathland percentage coverage and topsoil organic carbon content are the most important variables. Whilst for Oystercatcher, grassland and heathland percentage cover are almost equivalent in their importance followed by longitude. For Redshank and Snipe, available water capacity and heathland are the most important covariates in predicting abundance.

### Breeding wader population estimates from 2002 to 2019

The random forest regression model was used to predict species abundance over *all* (n=3992) Shetland BBS 1km squares. The model gave a mean estimate together with lower and upper confidence intervals (5% and 95% percentiles respectively), across every year the survey was run (2002 to 2019). The population estimate results are shown in Appendix ?? and plotted in Figure 28

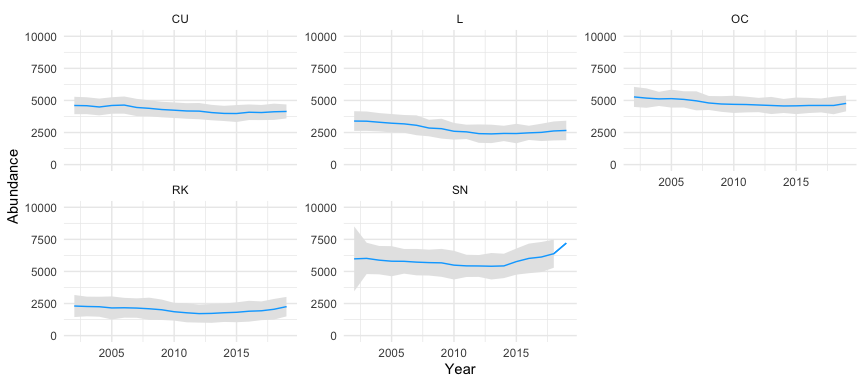


Figure 28: Breeding wader population estimates - number of pairs of breeding waders by species, 2002 to 2019. Grey shaded area indicates upper and lower confidence intervals

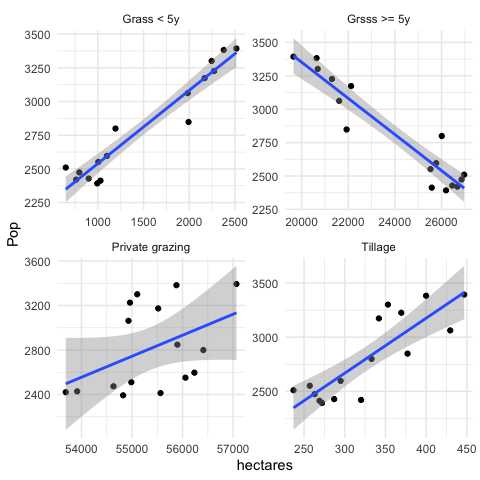
Across the years 2002 to 2019 the abundance of breeding wader pairs across all species appear to have decreased, with the exception of Snipe. The most significant decline was for Lapwing. Note that the confidence intervals for Snipe are highly variable in certain years. Table 13 shows the change in breeding wader abundance by species between 2002 and 2019.

Table 13: Change in breeding wader abundance across all species, between the years 2002 and 2019

|  |  |  |  |
| --- | --- | --- | --- |
| Species | 2002 | 2019 | % Change |
| CU | 4597 | 4088 | -11.1 |
| L | 3474 | 2638 | -24.1 |
| OC | 5269 | 4760 | -9.7 |
| RK | 2390 | 2248 | -5.9 |
| SN | 6043 | 7391 | 22.3 |

### Lapwing abundance association with grassland holdings

Data from 2002 to 2017 for grassland holdings in hectares categorised as: tillage, grassland less than five years old (reseeded grassland), grassland five years or greater and private grazing were tested for normality, and found to be log-normally distributed. A linear regression was plotted for Lapwing abundance versus each grassland categorisation (in Hectares) as shown in Figure 29.

 Figure 29 shows that Lapwing population size is positively associated with the size of grassland that is less than five years old (that grassland which has been reseeded) and tillage (land prepared for spring crops). Figure 29 also shows that Lapwing population abundance is negatively associated with grassland that is not reseeded (Grass >= 5y).

The r package gamlss (Rigby and Stasinopoulos [2005](#ref-gamlss)) was used to fit a log-normal model for Lapwing population abundance with grassland less than 5 years old as a covariate. The grassland < 5 years old was a statistically significant covariate (p<0.0001), and the residuals of the model were approximately normally distributed as shown in Figure 30.

## GAMLSS-RS iteration 1: Global Deviance = 192.9757   
## GAMLSS-RS iteration 2: Global Deviance = 192.9757

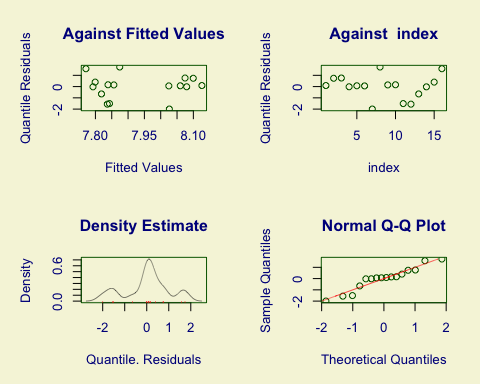


Figure 30: Residuals for model fit of Lapwing population estimates to Shetland improved grassland less than five years old. Agricultural statistics taken from Scottish Agricultural Survey

## Spatial abundance distriution of breeding waders

The population estimates results in 3.8.4 are based on a spatial prediction. As such they can be plotted to show how species abundance is spatially distributed across Shetland. Figure 31 shows the abundance distribution, per 1km (density) for each species in 2019.

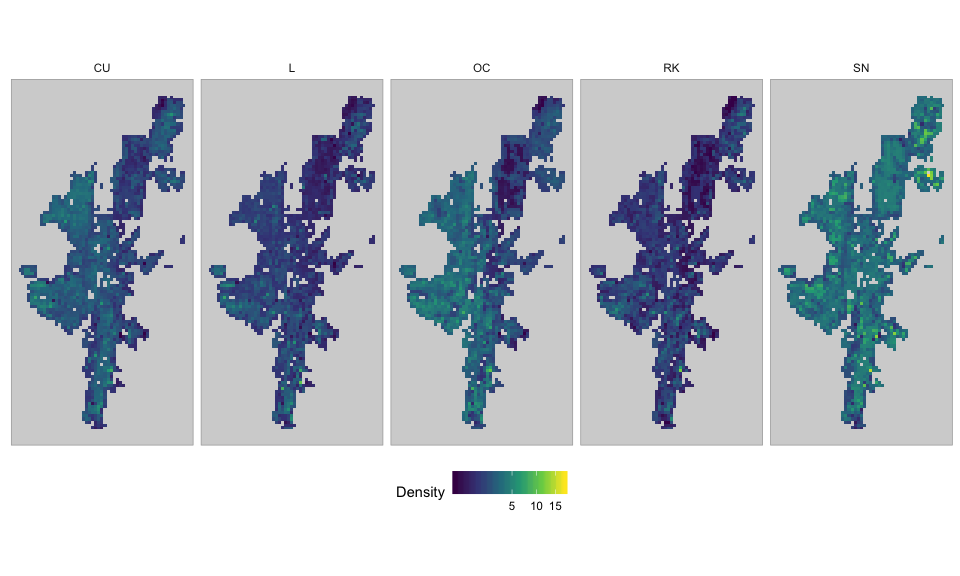


Figure 31: Spatial abundance distribution of breeding waders for 2019

It can be seen that Snipe are widespread, and that Curlew and Oystercatcher are significantly present on the western side of Shetland, as suggested by the variable importance plots in Appendix ??. Lapwing and Redshank have the lowest population densities and seem to be concentrated in the south west of Shetland.

### Net spaital abundance change by species, between 2002 and 2019

Given the spatial abundance distribution for 2002 and 2019, it is possible to plot the net change in breeding wader abundance, for each 1km, between the two years. This is shown in Figure 32.

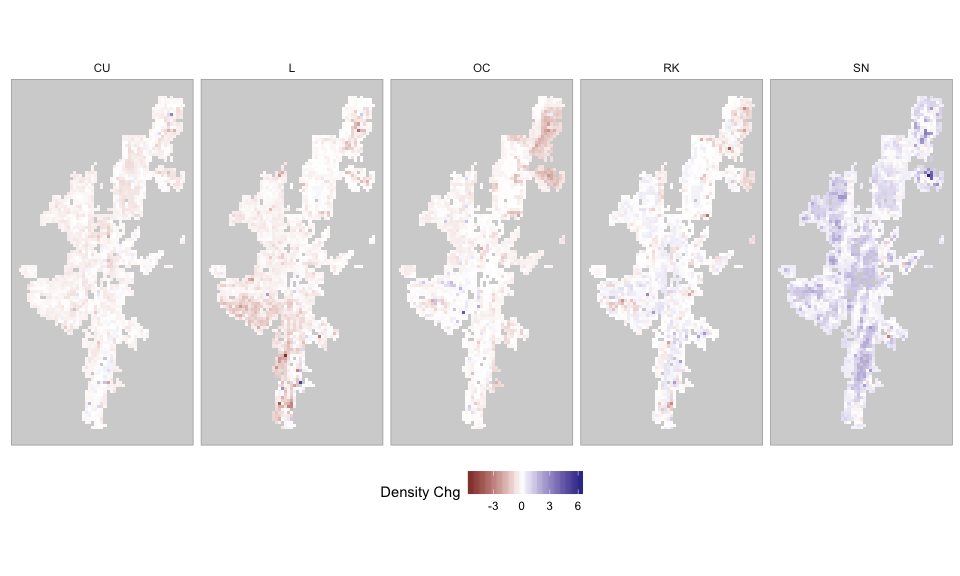


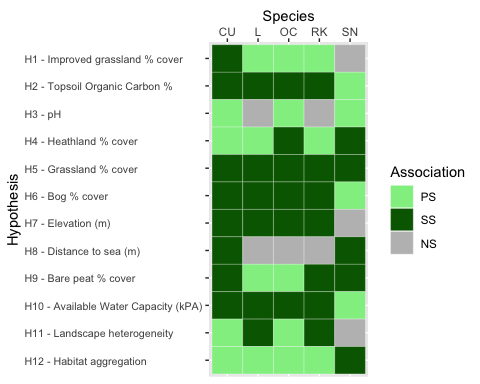
Figure 32: Change in breeding wader density (count/km2) between 2002 and 2019 on Shetland

It can be seen that the drop in abundance as shown in Appendix ?? is reflected in the net change plots in Figure 32.

The initial analysis of the status of surveyed 1km squares (see Figure 7) suggested that Lapwing had undergone a decline, and Figure 32 shows that this has occurred in the south west and north east over the two analysis periods. Oystercatcher appear to have undergone a significant decline in the north east of Shetland (the islands of Unst and Fetlar). Where as Snipe appear to have increased in most upland areas of Shetland, and in particular on the Fetlar RSPB reserve.

# Discussion

The results indicate there are a number of statistically significant associations between breeding wader abundance and the various covariates outlined in the hypotheses in the introduction in Section 1.6. Figure 33 summarises the results against each of these hypotheses. If the results were in agreement across both analysis periods they are marked as strongly supported (SS) and if only in one period partially support (PS). If the hypothesis is not supported by the results at all, or it is not statistically significant, then it is marked as not supported (NS).

 ## Discussion of evidence for each hypothesis

### Improved grassland coverage (H1)

In the 1980-90s agricultural subsidies encouraged the farming of high numbers of upland sheep (Shetland Islands Council [2016](#ref-Shetland_Islands_Council2016-tz)), which in turn supported the “improvement” of marginal in-bye grassland so as to produce silage as a winter feed for sheep. Improved fields were also treated with lime and fertiliser so as to artificially boost yields of grass for silage. The practise of rotating re-seeded improved grassland and a spring-sown crop is not widely undertaken in Shetland any longer. The area of ground in cultivation for crops (excluding improved grassland) in Shetland fell by 90% between 1971 and 2008 (Shetland Islands Council [2016](#ref-Shetland_Islands_Council2016-tz)). This change in agricultural practice has meant that the majority of grassland sward has become permanent, and due to ongoing sheep grazing is therefore likely to be uniform in structure. The practise of annual fertilising and liming will increase pH which is good for invertebrate abundance, but fast growing and uniform sward, ready to cut in the early summer before the end of the wader nesting season, may not be ideal for breeding waders.

The 2002-2010 analysis period for improved grassland is positively associated with Curlew and Redshank, and negatively associated with Snipe. The second analysis period from 2011-2019 has improved grassland positively associated with all wader species, apart from Snipe which is again negatively associated. The second analysis period validates the hypothesis that there is a positive association between breeding wader density and improved grassland, whilst the first analysis period partially supports this. Snipe are known to breed in dense vegetation, marshes and bogs (Robinson [2005](#ref-Robinson2005-qa)) and the results agree with this, in that their abundance response is negatively associated with improved grassland in both analysis periods.

For Lapwing and Oystercatcher, the population change ratio between 2002 and 2019 in Figure ?? was positively associated with increased improved grassland coverage and this agrees with the literature (Dallimer et al. [2010](#ref-Dallimer2010-kg); Bell and Calladine [2017](#ref-Bell2017-iy)), in that population densities appear stable or increasing on low intensity improved grassland. However Figure 32 clearly shows that between 2002 and 2019 Lapwing abundance have declined significantly in the south west mainland, where improved grassland is mostly located. The results in Section 3.8.5 suggest that Lapwing populations are strongly associated with the overall size of reseeded improved grassland in Shetland. The reduction in sheep headage in Shetland (Shetland Islands Council [2016](#ref-Shetland_Islands_Council2016-tz)) could be driving the reduced acreage of reseeded improved grassland over this period. The reduction in spring-sown crops and improved grassland together with lower natural food availability in the upland habitat of Shetland, could have driven the overall declining Oystercatcher and Lapwing populations to increased densities in the remaining improved grassland habitat, due to greater availability of food to feed chicks. Mccallum et al observed this phenomenon in their paper:

*“Lapwing distribution is being constrained between intensively managed lowland farmland with favourable soil conditions and upland sites where lower management intensity favours Lapwings but edaphic conditions limit their distribution”* (McCallum et al. [2018](#ref-McCallum2018-gx))

Current agricultural practices on the more intensively-managed improved grasslands may be incompatible with successful Lapwing and Oystercatcher breeding. For example, fast growing grass intended for silage may not provide the open view that Lapwing require throughout their nesting period, which could lead to nest abandonment.

### Topsoil organic carbon content (H2)

Peat rich soils have an organic carbon content >40% by weight (McCallum et al. [2016](#ref-McCallum2016-jt)). Figure 17 shows that this is the case for the majority of Shetland soils. A higher organic content is associated with greater soil acidity, which in turn leads to a lower earthworm abundance (McCallum et al. [2018](#ref-McCallum2018-gx)) than in less acidic soils.

Both analysis periods from 2002 to 2019 have strong support for the hypothesis that topsoil organic carbon content is negatively associated with breeding wader abundance. The only exception is for Snipe during 2011-2019 where there was no statistically significant association. In terms of the population change ratio, Snipe population change is positively associated with an increase in topsoil organic carbon content, and Lapwing are not. This is consistent with the literature in that Snipe prefer wetlands which are likely to be acidic bogs that have high organic carbon content, and Lapwing prefer bare fields or short sward grassland that will have relatively low organic carbon content in the topsoil.

### pH (H3)

Soil pH together with topsoil organic carbon content, is a key indicator of soil health (McCallum et al. [2016](#ref-McCallum2016-jt), [2018](#ref-McCallum2018-gx)) as measured by earthworm density, which falls significantly if the soil pH is less than 5.5. As Shetland soils are mostly acidic peatland, especially in upland areas, a significant proportion of soils have a pH below this level (see Figure 17). It is therefore likely that peatland areas are negatively associated with earthworm abundance. Improved grassland is typically treated with lime and fertiliser in-order to boost growth, which has the side effect of increasing the pH which in-turn increases earthworm abundance. As such a higher pH would be expected to give a positive association with breeding waders.

The 2002-2010 analysis period for pH has no statistically significant associations with pH, apart from Snipe, which were positively associated. The 2011-2019 analysis period had positive associations for Curlew and Oystercatcher. So it seems that the second analysis period partially supports hypothesis *H3*, but the 2002-2011 analysis period does not. In terms of the population change ratio, pH is positively associated with Lapwing and Oystercatcher population change, and negatively associated with Snipe which is consistent with the literature. It is possible that Lapwing and Oystercatcher are again nesting in marginal land associated with higher pH, such as rough grassland. Snipe prefer wetlands that are likely to have significant rush vegetation which is typical for a relatively low pH soil.

### Heathland percentage coverage (H4)

Farmland in Shetland is currently neither highly intensive nor extensive, and primarily occurs as improved grassland for the purposes of winter animal feed. Wader feeding habitat such as improved grassland or shallow wetland is typically embedded within a mosaic of semi-natural blanket bog, unimproved grassland and heath. Heathland habitat within 500m to 2500m (**???**) too wetland or improved grassland, makes it ideal for nesting waders such as Curlew, Redshank and Snipe. This is because they can exploit the availability of earthworms and invertebrates too feed their chicks whilst nesting in longer heathland vegetation that gives their nest sites greater protection from predators.

The 2002-2010 analysis period partially supports the hypothesis that heathland is positively associated with all wader species (Figure 20), with only Oystercatcher and Snipe having a positive association with increased heath percentage coverage. This is expected for Snipe, but Oystercatcher are expected to nest on improved or rough grassland (Bell and Calladine [2017](#ref-Bell2017-iy)). The second analysis period from 2011-2019 shows that all breeding waders have a positive association with increased heath coverage per . Thess results together with the result of the population change ratio association for heathland percentage coverage (see Figure 22), suggests that non-heathland species such as Lapwing and Redshank are being displaced from their preferred habitats of improved grassland and wetland respectively into heathland, due to changes in habitat over the two analysis periods.

Anecdotal evidence suggests the condition of heathland habitats in Shetland since the late 1990s has improved, due to the disappearance of sheep headage payments that supported very high numbers of sheep in the hills, and the initiation of agri-environment scheme funding for removal of sheep from the uplands where heathland is significant. Greater vegetation cover and low disturbance levels during the breeding season is possibly providing more suitable nesting conditions in the second survey period, leading to a universal positive association between heathland habitat and breeding waders of all species.

### Grassland percentage coverage (H5)

Within the EUNIS landcover categorisation dataset, *grassland* contains improved and semi-natural grasslands and so is quite a broad class of habitat (EUNIS [2019](#ref-Eunis2019-dy)). It was expected that grassland would have a strong association with all breeding wader species due to the presence of both feeding (earthworms and other invertebrates) and nesting sites relatively well hidden in longer sward, and this is the case across both analysis periods. Perhaps more interesting is that Redshank are the only species to have a statistically significant association in the population change model for grassland association. Breeding Redshank population change is negatively associated with grassland percentage coverage (see Figure 22).

The reason for this could be that Redshank require a structurally diverse sward, with areas of tall vegetation to hide their nest, and patches of open vegetation and shallow pools for chicks to feed in (Sharps et al. [2016](#ref-Sharps2016-jd)). The maintenance of such conditions relies on grazing practices to create such vegetation diversity, whilst avoiding nest trampling during the breeding season. From an agricultural point of view, suitable redshank habitats are often of poor value and if they have not been agriculturally improved with drainage, they may have become under-grazed leading to habitat which is not suitable for Redshank (Farm Advisory Service [2017](#ref-Farm_Advisory_Service2017-xu)). Agri-environment schemes have supported the conservation of wetland and wet grassland, primarily through the exclusion of livestock during the breeding season with a light grazing period in the late summer and autumn to remove the annual growth of vegetation. However, if the weather and ground conditions are unsuitable at the end of the year for livestock grazing, these wetlands may be left ungrazed, creating a dense sward of low value as breeding habitat in the following spring.

### Bog percentage coverage (H6)

A significant amount of Shetland is covered in peatland. Healthy peat is relatively acidic due to its permanently water logged state, and so there is unlikely to be a sufficient earthworm supply to support breeding wader populations. The results suggest that this is the case with both analysis periods showing negative association between breeding wader abundance and bog percentage coverage, with the exception of Snipe where association was not statistically significant in the first analysis period. In terms of the population change ratio, it appears that Redshank are positively associated with increased bog percentage coverage (see Figure 22). This could be because Redshank prefer a mosaic of wet habitat and grassland when breeding, or it could be that grassland has become less favourable in terms of food availability and so Redshank are nesting in bog habitat as the next best alternative. The population change ratio for Curlew is negatively associated with increased bog percentage landcoverage. This presumably is because nesting Curlew prefer tall vegetation typical of rough grassland and heathland (Franks et al. [2017](#ref-Franks2017-co)).

### Elevation (H7)

Habitat at higher elevations within Shetland is typically heathland, exposed rock or bog, and heavily exposed to the extremes of a marine sub-arctic climate. As such food supply for breeding waders will be relatively low compared to lowland areas. This hypothesis (H7) is confirmed by the results, which show there is a strongly supported negative association between breeding wader abundance and elevation, except for Snipe, which was not statistically significant (see Figure 20). The population change model for elevation has only statistically significant associations for Curlew and Redshank (see Figure 22). The fact that breeding Curlew abundance is negatively associated with elevation could again be due to shorter nesting vegetation and availability of earthworms at higher elevations, which are typically peat bog or rock. The Redshank population change is positively associated with bog and heath cover which tend to be significant habitat at higher elevations, and negatively associated with grassland cover. Whilst lowland habitats are less suitable, higher altitude habitats may have become more attractive with reduced sheep grazing.

### Distance from the sea (H8)

The results for the distance from the sea to the nest site vary by significantly by species. In the first analysis period Curlew, Oystercatcher and Snipe have positive associations for distance from the sea to nest locations, whilst Lapwing have a negative association and Redshank is not statistically significant. The second analysis period has the same results apart from Oystercatcher which now have a negative association. Those species that show greater nesting abundance at large distances from the coast may be responding to pressure from coastal predators. Shetland has a significant coastal population of Eurasian Otter *Lutra lutra* (JNCC [2020](#ref-Jncc2020-tp)) and gulls such as the Great black backed gull *Larus marinus* that may predate wader eggs or chicks. For species that respond negatively to the distance from the sea, Lapwing and Oystercatcher, this may be because their preference is for improved grassland that is predominantly in the south mainland and typically only a few hundred meters from the sea. Snipe breeds in moorland habitats that predominate in at higher elevations that are typically further away from the coast.

The only statistically significant results for the population change model are for Curlew and Oystercatcher, which both show a negative association between population change over the two analysis periods and distance from the sea. This maybe due to changes within inland nesting territories, such as the decline in available food resources over time.

### Bare peat percentage coverage (H9)

Shetland has a significant amount of degraded peatland, and a proportion of this is degraded to the point where the thin vegetative top layer of sphagna known as the *acrotelm* has completely died, leaving the deep acidic and water logged peat store known as the *catotelm* exposed. The exposed bare peat layer then further degrades overtime. The results for the two analysis periods show that as expected wader breeding abundance is either strongly or partially supported by a negative association with bare peat percentage landcover. This is most likely due to their being very little food resources available and suitable nesting opportunities in such habitat. Although the population change model results show that Redshank are positively associated with this covariate. This could be because bare peatland forms part of upland mosaics that include wetlands and heath, where Redshank are known to nest.

### Available water capacity (H10)

Heathland together with unimproved grassland and blanket bog characterises the majority of habitat within the Shetland archipelago. Breeding wader densities in these semi-natural mosaic are found to be consistently higher when surrounding landscapes have more wetland at a proximity ranging from 500m to 2500m, indicating the importance of wetland availability (Jóhannesdóttir et al. [2019](#ref-Johannesdottir2019-uh)). Both analysis periods show that all breeding wader abundance, apart from Snipe, is negatively associated with increasing water capacity of the topsoil. This is contrast to the hypothesis made in the introduction in Section ?? that postulated that wader abundance would increase as available water capacity increases. It appears that whilst all waders have a preference for damp or wet habitat, only Snipe prefer true water logged habitat that is characteristic of the upland wetlands. Although it appears that the population change model shows that Oystercatcher and Redshank are also tending to show a positive association in terms of population change and available water capacity (see Figure 22). This association could explain why there have been significant declines of Oystercatcher and Redshank between 2002 and 2019 in the north east of Shetland on the islands of Unst and Fetlar (see Figure 32). Here the available water capacity (AWC) in the soil is relatively low (see Appendix ??), due to the serpentine geology in the north east of Shetland. So if the climate is becomming increasingly dry during the breeding season this could be forcing both Oystercather and Redshank too areas with higher AWC, due to the lack of topsoil invertebrates, and hence the decline of these two species from this area of Shetland.

### Landcover heterogeneity (H11)

Marginal entropy is a metric to measure the thematic complexity, or heterogeneity of landcover. As postulated in (**???**) landscapes that have relatively low agricultural productivity and that are farmed at low intensity and extent, often maximise species diversity and resources. As Shetland is characterised by these type of agriculture practices, which in turn are embedded within a semi-natural mosaic of wetland, heath, rough grassland and peat bogs, it is expected that wader breeding abundance is positively associated with increased marginal entropy. There appears to be strong support for this hypothesis across all wader species, except Snipe, in the second analysis period (see Figure 26. Snipe are negatively associated with increasing marginal entropy and this is likely due to their preference (Farm Advisory Service [2017](#ref-Farm_Advisory_Service2017-xu)) for aggregated wetland and moorland habitat when nesting. This exists extensively in Shetland.

### Landcover homogeneity (H12)

Relative mutual information seeks to measure the degree of spatial aggregation of habitat classes within the wider landscape it is situated in. It is expected that breeding waders prefer a mosaic of distinct habitats that allows chicks to walk between nesting and feeding habitats, as demonstrated in Section 4.0.11. The hypothesis that waders would negatively respond to increased spatial aggregation of habitat, is not supported by the results. The results suggest a positive association between all wader species and relative mutual information in the first analysis period, and no statistically significant association apart from Snipe, that are positively associated, in the second analysis period. It is expected that Snipe would be positively associated with increasing relative mutual information as they are specialists of aggregated wetland and moorland habitats. Despite the non-conclusive evidence for the two individual analysis periods, the population change model does have statistically significant negative associations for increasing relative mutual information and Lapwing and Oystercatcher abundance changes. This suggests that both these species have shown lower nesting densities when habitat homogeneity increases. A question of scale (1km)???? Look at negatiev skew in associations.

Switch in aggregated habitat too more diverse habitat over time - due to less intensive sheep grazing?

## Shetland wader population trends and implications for conservation outcomes

The results from the Shetland wader population trends show that breeding waders across Shetland have declined, but not nearly as much as those on the UK mainland. The 2015 BTO Breeding Bird Survey (Harris, S.J., Massimino, D., Newson, S E., Eaton, M.A., Marchant, J.H., Balmer, D.E., Noble, D.G., Gillings, S., Procter, D. & Pearce-Higgins, J.W. [2016](#ref-Harris_SJ2016-yw)) estimates that across Britain farmland wader breeding populations between 1995 and 2015 declined -19% for Oystercatcher, -43% for Lapwing, -48% for Curlew, -39% for Redshank.

A 1993 RSPB wader survey [ref] was conducted inorder to estimate the overall numbers of breeding waders on *lowland* Shetland (a total area of 398 km ). In 1986 a survey was undertaken of Shetland *moorland* (ref), a total area of 1,019km. Although both surveys are separated by some nine years, the sum of both surveys gives an estimate for the baseline for the results of this study. Table ?? shows the results of the combine 1986 and 1993 surveys, together with the population estimates for the BBS surveys in 2002 and 2019.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Species | 1986 | 1993 | 1986 + 1993 | 2002 | 2019 |
| CU | 1950 | 2025 | 3975 | 4567 | 4088 |
| L | 650 | 1950 | 2600 | 3474 | 2638 |
| OC | 1700 | 3300 | 5000 | 5269 | 4760 |
| RK | 110 | 1075 | 1275 | 2390 | 2248 |
| SN | 800 | 2650 | 3450 | 6043 | 7391 |

It can be seen that for Curlew, Lapwing and Oystercatcher 2019 and 1986/1993 population estimates are broadly similar, but there was a noticeable increase in Redshank and Snipe. For Snipe it is possible that the methodology for estimating their numbers has improved, such that the 1986/1993 survey underestimated their numbers.

Comparing 2019 to 2002 it does look as if there was a general decline in each species, apart from in Snipe. The most significant decline being Lapwing the analysis shown in Figure xxx shows that Lapwing decline are strongly associated with a reduction in reseeding of improved grassland for agricultural silage. The clear association between declines of lapwing and temporary grassland is statistically significant (p<0.0001) and it suggests that the 24% drop in the Shetland Lapwing population could be due to the significant decline of bare spring-sown grassland and tillage. This finding is also suggested as the major cause of Lapwing declines in a 25 year study of a population of farmland breeding waders (Bell and Calladine [2017](#ref-Bell2017-iy)). Lapwing appear to benefit from agricultural practices that create bare earth or a very short sward (<5cm) every spring, or more generally nesting sites with an open view. It appears that practices over time, driven by subsidies, have created a grass monoculture of low value to lapwings. This evidence could help in shaping agri-environment schemes, where focus to-date focused has been on improving nesting habitats, rather than focusing on ensuring feeding habitats are maintained or improved. Ideally a mosaic of both habitats would be created under these schemes. This result suggests that low intensity agriculture is a critical factor in the success of breeding wader populations, and how farmers can be important stewards of nature.

Something about Snipe response here. Variability of estimates due to sampling method for drumming snipe?

The most recent estimates of the Shetland breeding population are from surveys carried out in the late 1990s: 2,300 pairs of Curlew, 1,740 pairs of Lapwing, 3,350 pairs of Oystercatcher, 1,170 pairs of Redshank and 3,450 pairs of Snipe (Harvey et al. 2004)

## The significance of soil health for breeding wader populations

An important area that has arisen through analysing wader densities and population change is the interaction between semi-natural and improved grassland habitats, and how this interaction drives food availability for wader chicks. Breeding wader research has recently started to focus on understanding the importance of soil health (McCallum et al. [2018](#ref-McCallum2018-gx), [2016](#ref-McCallum2016-jt)) in creating adequate food supplies for breeding waders. Understanding earthworm and invertebrate density gradients in different habitat types, and any interaction effects between environmental covariates would be beneficial to improving wader populations and is a potential area of future work. There has been some related work (Pansu et al. [2015](#ref-Pansu2015-ad)) whereby environmental-DNA of soil samples was sequenced to make landscape-scale assessments of soil-biodiversity and its drivers.

As the population change model for Lapwing clearly shows (ref), their population whilst declining, has increased in density on lowland improved grassland cover. The Agricultural Census data(ref) show a decline in this habitat type. This suggests that lack of available feeding habitat could be driving this particular change in spatial distribution. One way of improving feeding habitat in areas where suitable lowland habitat is not available, could be through the application of lime on upland where there is a greater amount of suitable nesting habitats, but where food abundance is not sufficient due to the quality of the soil structure.

## Covariates that were not explored and future work

Vegetation structure - tall dense vegetation containing rushes is not preferred by Lapwing or Oystercatcher, as they require vegetation that gives them a clear line of sight for predators. when nesting. However dense and taller vegetation structure is ideal for waders in general as Redshank, Curlew and Snipe can hide their nests in the longer grass and rushes. Recent work (Howison et al. [2018](#ref-Howison2018-gc)) has focused on developing radar‐based remote sensing technique (Sentinel‐1 C‐band radar) for large‐scale quantification of agricultural land‐use intensity and vegetation, across human‐dominated landscapes.

Predation - there is significant potential for wader nest and chick predation on Shetland. Possible predators include hedgehogs, rats, corvids and coastal predators including otter and gulls. It was not possible to obtain relevant survey data in order to test this hypothesis and remains an important area of possible future survey work and data analysis.

Weather - changes in weather patterns might also have a significant long term effect on breeding waders on Shetland. For example very wet spring weather may mean it is not posssible for grassland to be reseeded or spring crops, such as barley, to be sown. As discussed previously, this is a key habitat for Lapwing and Oystercatcher. An important piece of future of work would be to examine the impact of future scenarios of climate change. If the wader breeding season has more severe weather, this may significantly affect wader food sources or productivity in general.

## Conclusion

Main conclusions:

Shetland is, relatively speaking, a strong-hold for Farmland waders.

* driven the overall declining Oystercatcher and Lapwing populations to increased densities in the remaining improved grassland habitat, due to greater availability of food to feed chicks. the importance of feeding habitat, agri-schemes that have sprinfg sown crops or spring tillage
* suggests that non-heathland species such as Lapwing and Redshank are being displaced from their preferred habitats of improved grassland and wetland respectively into heathland
* redshank are being displaced from grasslands over time due to a lack of heavy grazung at the right time - it could be that grassland has become less favourable in terms of food availability and so Redshank are nesting in bog habitat as the next best alternative. heathland, elevation and bog coverage are all positive associations over time.
* if the climate is becomming increasingly dry during the breeding season this could be forcing both Oystercather and Redshank too areas with higher AWC, due to the lack of topsoil invertebrates, and hence the decline of these two species from this area of Shetland.

# References

Abdi, Abdulhakim Mohamed. 2020. “Land Cover and Land Use Classification Performance of Machine Learning Algorithms in a Boreal Landscape Using Sentinel-2 Data.” *GISci. Remote Sens.* 57 (1): 1–20.

Andy Liaw and Matthew Wiener. 2002. “Classification and Regression by randomForest.” *R News* 2/3 (December): 18–22.

Bell, Michael V, and John Calladine. 2017. “The Decline of a Population of Farmland Breeding Waders: A Twenty-Five-Year Case Study.” *Null* 64 (2): 264–73.

BirdLife International. 2015a. “Gallinago Gallinago. The IUCN Red List of Threatened Species 2015: e.T22693097A66707592.” <https://www.iucnredlist.org/species/22693097/66707592>.

———. 2015b. “Numenius Arquata. The IUCN Red List of Threatened Species 2015: e.T22693190A60033961.” <https://www.iucnredlist.org/species/22693190/60033961>.

———. 2015c. “Tringa Totanus. The IUCN Red List of Threatened Species 2015: e.T22693211A60036510.” <https://www.iucnredlist.org/species/22693211/60036510>.

———. 2017. “Vanellus Vanellus (Amended Version of 2016 Assessment). The IUCN Red List of Threatened Species 2017: e.T22693949A111044786.” <https://dx.doi.org/10.2305/IUCN.UK.2017-1.RLTS.T22693949A111044786.en.>

———. 2019. “Haematopus Ostralegus. The IUCN Red List of Threatened Species 2019: e.T22693613A154998347.” <https://dx.doi.org/10.2305/IUCN.UK.2019-3.RLTS.T22693613A154998347.en>.

Blake, Duncan. 2020. “Scotland Bare Peat.” https://github.com/duncansnh/Bare-peat.

Broek, J van den. 1995. “A Score Test for Zero Inflation in a Poisson Distribution.” *Biometrics* 51 (2): 738–43.

Brown, Daniel, Jeremy Wilson, David Douglas, Patrick Thompson, Simon Foster, Neil Mc Culloch, James Phillips, et al. 2015. “The Eurasian Curlew – the Most Pressing Bird Conservation Priority in the UK?” *British Birds* 108 (November).

Butchart, Stuart H M, Matt Walpole, Ben Collen, Arco van Strien, Jörn P W Scharlemann, Rosamunde E A Almond, Jonathan E M Baillie, et al. 2010. “Global Biodiversity: Indicators of Recent Declines.” *Science* 328 (5982): 1164–8.

Dallimer, Martin, Lorenzo Marini, Andrew M J Skinner, Nick Hanley, Paul R Armsworth, and Kevin J Gaston. 2010. “Agricultural Land-Use in the Surrounding Landscape Affects Moorland Bird Diversity.” *Agriculture, Ecosystems and Environment* 139 (4): 578–83.

Davison, A. C., and D. V. Hinkley. 1997. *Bootstrap Methods and Their Applications*. Cambridge: Cambridge University Press. <http://statwww.epfl.ch/davison/BMA/>.

Donnelly, D and Buckler, C. 2014. “Topsoil pH Map (Median).”

ESA. 2020. “User Guides - Sentinel-2 MSI - Sentinel Online.” <https://sentinel.esa.int/web/sentinel/user-guides/sentinel-2-msi>.

EUNIS. 2019. “EUNIS Land Cover Scotland - Data.gov.uk.” <https://data.gov.uk/dataset/0acb95af-f367-4408-be87-f862651a6b16/eunis-land-cover-scotland>.

Eze, Samuel, Sheila M Palmer, and Pippa J Chapman. 2018. “Soil Organic Carbon Stock in Grasslands: Effects of Inorganic Fertilizers, Liming and Grazing in Different Climate Settings.” *J. Environ. Manage.* 223 (October): 74–84.

Farm Advisory Service. 2017. “Management and Conservation for Farmland Waders.” TECHNICAL NOTE TN688. FAS.

Fiske, Ian, and Richard Chandler. 2011. “Unmarked: An R Package for Fitting Hierarchical Models of Wildlife Occurrence and Abundance.” *Journal of Statistical Software, Articles* 43 (10): 1–23.

Franks, Samantha E, David J T Douglas, Simon Gillings, and James W Pearce-Higgins. 2017. “Environmental Correlates of Breeding Abundance and Population Change of Eurasian Curlew Numenius Arquata in Britain.” *Bird Study* 64 (3): 393–409.

Gagkas, Z., Lilly, A., Baggaley, N. & Donnelly, D. 2019. “Map of Available Water Capacity of Soils in Scotland.”

Greenwell, Brandon, Brad Boehmke, and Bernie Gray. 2020. *Vip: Variable Importance Plots*. <https://CRAN.R-project.org/package=vip>.

Harris, S.J., Massimino, D., Newson, S E., Eaton, M.A., Marchant, J.H., Balmer, D.E., Noble, D.G., Gillings, S., Procter, D. & Pearce-Higgins, J.W. 2016. “The Breeding Bird Survey 2015 BTO Research Report 687.” 687.

Harris, S J, D Massimino, S Gillings, M A Eaton, D G Noble, D E Balmer, D Procter, J W PearceHiggins, and P Woodcock. 2018. “The Breeding Bird Survey 2017.” BTO/JNCC/RSPB.

Hesselbarth, Maximilian H. K., Marco Sciaini, Kimberly A. With, Kerstin Wiegand, and Jakub Nowosad. 2019. “Landscapemetrics: An Open-Source R Tool to Calculate Landscape Metrics.” *Ecography* 42: 1648–57.

Hijmans, Robert J. 2020. *Raster: Geographic Data Analysis and Modeling*. <https://CRAN.R-project.org/package=raster>.

Hollister, Jeffrey, and Tarak Shah. 2017. *Elevatr: Access Elevation Data from Various Apis*. <http://github.com/usepa/elevatr>.

Howison, Ruth A, Theunis Piersma, Rosemarie Kentie, Jos C E W Hooijmeijer, and Han Olff. 2018. “Quantifying Landscape-Level Land-Use Intensity Patterns Through Radar-Based Remote Sensing.” Edited by Tomas Pärt. *J. Appl. Ecol.* 55 (3): 1276–87.

JNCC. 2020. “Otter (Lutra Lutra).” <https://sac.jncc.gov.uk/species/S1355/>.

Jóhannesdóttir, Lilja, Jennifer A Gill, José A Alves, Sigmundur H Brink, Ólafur Arnalds, Verónica Méndez, and Tómas Grétar Gunnarsson. 2019. “Interacting Effects of Agriculture and Landscape on Breeding Wader Populations.” *Agric. Ecosyst. Environ.* 272 (February): 246–53.

Kuhn, Max, and Hadley Wickham. 2020. *Tidymodels: Easily Install and Load the ’Tidymodels’ Packages*. <https://CRAN.R-project.org/package=tidymodels>.

Lilly, A, Baggaley, N and Donnelly, D. 2012. “Map of Soil Organic Carbon in Top Soils of Scotland.”

Macdonald, Benedict. 2019. *Rebirding - Rewilding Britain and Its Birds*. Pelagic.

McCallum, Heather M, Kirsty J Park, Mark G O’brien, Alessandro Gimona, Laura Poggio, and Jeremy D Wilson. 2018. “Soil pH and Organic Matter Content Add Explanatory Power to Northern Lapwing Vanellus Vanellus Distribution Models and Suggest Soil Amendment as a Conservation Measure on Upland Farmland.” *Ibis*, June, 677–87.

McCallum, Heather M, Jeremy D Wilson, David Beaumont, Rob Sheldon, Mark G O’Brien, and Kirsty J Park. 2016. “A Role for Liming as a Conservation Intervention? Earthworm Abundance Is Associated with Higher Soil pH and Foraging Activity of a Threatened Shorebird in Upland Grasslands.” *Agric. Ecosyst. Environ.* 223 (May): 182–89.

Melbourne, Brett A, and Alan Hastings. 2008. “Extinction Risk Depends Strongly on Factors Contributing to Stochasticity.” *Nature* 454 (7200): 100–103.

Newson, Stuart E, Karl L Evans, David G Noble, Jeremy J D Greenwood, and Kevin J Gaston. 2008. “Use of Distance Sampling to Improve Estimates of National Population Sizes for Common and Widespread Breeding Birds in the UK.” *J. Appl. Ecol.* 45 (5): 1330–8.

Nowosad, Jakub, and Tomasz F Stepinski. 2019. “Information Theory as a Consistent Framework for Quantification and Classification of Landscape Patterns.” *Landsc. Ecol.* 34 (9): 2091–2101.

Oesterle, M, and R Wildmann. 2003. “Land Parcel Identification as Part of the Integrated Administation and Control System (Iacs).”

Pansu, Johan, Sébastien De Danieli, Jérémy Puissant, Jean-Maxime Gonzalez, Ludovic Gielly, Thomas Cordonnier, Lucie Zinger, et al. 2015. “Landscape-Scale Distribution Patterns of Earthworms Inferred from Soil DNA.” *Soil Biol. Biochem.* 83 (April): 100–105.

Paterson, Edward. 2011. “Geochemical Atlas for Scottish Topsoils.” *Macaulay Land Use Research Institute*.

Pebesma, Edzer. 2018. “Simple Features for R: Standardized Support for Spatial Vector Data.” *The R Journal* 10 (1): 439–46. <https://doi.org/10.32614/RJ-2018-009>.

Pettorelli, Nathalie, William F Laurance, Timothy G O’Brien, Martin Wegmann, Harini Nagendra, and Woody Turner. 2014. “Satellite Remote Sensing for Applied Ecologists: Opportunities and Challenges.” Edited by E J Milner-Gulland. *J. Appl. Ecol.* 51 (4): 839–48.

R Core Team. 2019. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.

Rigby, R. A., and D. M. Stasinopoulos. 2005. “Generalized Additive Models for Location, Scale and Shape,(with Discussion).” *Applied Statistics* 54.3: 507–54.

Robinson, R A. 2005. “Snipe Gallinago Gallinago.” <https://app.bto.org/birdfacts/results/bob5190.htm>.

Scottish Government, St Andrew’s House, Regent Road, and Edinburgh Eh1 3dg Tel: 2003. “Agriculture and Fisheries - Publications.” <https://www2.gov.scot/Topics/Statistics/Browse/Agriculture-Fisheries/PubScottishCensus>.

Sharps, Elwyn, Angus Garbutt, Jan G Hiddink, Jennifer Smart, and Martin W Skov. 2016. “Light Grazing of Saltmarshes Increases the Availability of Nest Sites for Common Redshank Tringa Totanus, but Reduces Their Quality.” *Agric. Ecosyst. Environ.* 221 (April): 71–78.

Shetland Islands Council. 2016. “European and External Relations Committee.”

Smart, J, A Amar, M O’Brien, P Grice, and K Smith. 2008. “Changing Land Management of Lowland Wet Grasslands of the UK: Impacts on Snipe Abundance and Habitat Quality.” *Anim. Conserv.* 11 (4): 339–51.

Smart, Jennifer, Jennifer A Gill, William J Sutherland, and Andrew R Watkinson. 2006. “Grassland-Breeding Waders: Identifying Key Habitat Requirements for Management.” *J. Appl. Ecol.* 43 (3): 454–63.

*The Birds of Shetland*. 2004. Helm County Avifaunas. Christopher Helm, pp. 204-205, 216-218, 235-236, 244-245; 247-248.

Thorup, O, ed. 2006. “Breeding Waders in Europe 2000.” International Wader Study Group.

Vickery, J A, J R Tallowin, R E Feber, E J Asteraki, P W Atkinson, R J Fuller, and V K Brown. 2001. “The Management of Lowland Neutral Grasslands in Britain: Effects of Agricultural Practices on Birds and Their Food Resources.” *J. Appl. Ecol.* 38 (3): 647–64.

Wilson, Andy M, Malcolm Ausden, and Tim P Milsom. 2004. “Changes in Breeding Wader Populations on Lowland Wet Grasslands in England and Wales: Causes and Potential Solutions.” *Ibis* 146 (November): 32–40.

Wood, S. N. 2003a. “Thin-Plate Regression Splines.” *Journal of the Royal Statistical Society (B)* 65 (1): 95–114.

———. 2003b. “Thin-Plate Regression Splines.” *Journal of the Royal Statistical Society (B)* 65 (1): 95–114.

Wright, Marvin N., and Andreas Ziegler. 2017. “ranger: A Fast Implementation of Random Forests for High Dimensional Data in C++ and R.” *Journal of Statistical Software* 77 (1): 1–17. <https://doi.org/10.18637/jss.v077.i01>.

Zuur, Alain F, Elena N Ieno, and Chris S Elphick. 2010. “A Protocol for Data Exploration to Avoid Common Statistical Problems.” *Methods Ecol. Evol.* 1 (1): 3–14.