University of Zurich UZH

Environmental factors affecting the nest site choice of the northern wheatear. A case study about Oenanthe oenanthe in Val Piora, Switzerland.

GEO 511 Master's Thesis

Author
Nadja Bernhard
12-715-868

Supervised by
Dr. Christoph Meier (christoph.meier@vogelwarte.ch)
Dr. phil. II Felix Liechti (felix.liechti@vogelwarte.ch)

Faculty representative
Prof. Dr. Robert Weibel

29.06.2019
Department of Geography, University of Zurich

Environmental factors affecting the nest site choice of the northern wheatear

A case study about *Oenanthe oenanthe* in Val Piora, Switzerland

A male northern wheatear (photo by Linus Cadotsch).

Nadja Bernhard

12-715-868

June 30, 2019

Prof. Dr. Robert Weibel

University of Zurich

Winterthurerstrasse 190

CH-8057 Zurich

robert.weibel@geo.uzh.ch

Dr. Christoph Meier

Swiss Ornithological Institute

Seerose 1

CH-6204 Sempach

christoph.meier@vogelwarte.ch

Dr. phil. II Felix Liechti

Swiss Ornithological Institute

Seerose 1

CH-6204 Sempach

felix.liechti@vogelwarte.ch

ABSTRACT

I intended to analyse the effect of environmental factors on the nest site choice of the northern wheatear (*Oenanthe oenanthe*) in Val the Piora, Switzerland. The study area contained 227 analysable nest sites and 241 randomly selected pseudo-absence points. For the analysis, I created the following variables using QGIS and R: elevation, slope, aspect, normalized difference vegetation index (NDVI), distance to the nearest path, cumulated distance to the three closest nests and terrain roughness. All these parameters were fitted in a generalized linear model (GLM), which estimated the likelihood of a nest at a given location. The resulting model implies a statistical preference for S to SW exposed hillsides around 1900 to 2100m a.s.l. as nest sites. Statistically, slopes with mostly uniform gradients of 15° to 30° appeared to be preferred. Moreover, a statistical preference could be established for even to slightly concave structures. Furthermore, my GLM might indicate that heterogeneous land cover with mostly sparse or no vegetation were favoured by northern wheatears despite the low availability of such unvegetated patches on the study area. The vicinity of other nest sites, i.e. small clusters, as well as the proximity to hiking paths were favoured.

The created semi-automatic algorithm extracted the snow coverage from digital camera pictures. The snow coverage was computed individually for 31 absence points and 38 presence points. The snow coverage was computed pixel-wise with an empirically defined RGB threshold value for snow. The proof of principle for snow coverage variables first snow-free day and last snow-covered day suggested a between-year variation of the snowmelt. Furthermore, the observed course of snow melt could be explained in accordance with the meteorological data. Finally, a GLM investigated the effects of the most influential environmental factors and the snow coverage on the nest site choice. The GLM suggests an absence of an effect of snow coverage on the nest site choice. Careful interpretation of these results are required because of the spatial, temporal and spectral limitations of the algorithm, the restrictions of the model and the early snowmelt in the Val Piora in comparison to the arrival of the northern wheatears.

This thesis supports the findings of the literature in slope, aspect, land cover and nest clustering for the Val Piora [Arlt and Part 2007; Wartmann 1985]. The nearest two nests were 15m apart in the Val Piora in 2013. This proximity largely undercuts distances between nests of other Alpine breeding grounds [Corti and Melcher 1958; Glutz von Blotzheim 1962]. However, the current data hinders the determination of the between-year variation of the dynamic variables snow coverage and land cover. Therefore, a more detailed environmental analysis of the nest site choice of the northern wheatear in the study area will be required.

Keywords: *Oenanthe oenanthe, northern wheatear, nest site selection, environmental factors, RGB, digital camera, snow coverage, DEM, slope, aspect, NDVI, distance to path, distance to nest, roughness, cluster, GLM*

ACKNOWLEDGEMENTS

Contents

1 INTRODUCTION

Northern wheatears occur almost worldwide, yet their Alpine and European population is not continuous [Knaus et al. 2018; Svensson et al. 2009]. Currently, this patchy distribution cannot be explained satisfyingly. The arguments explaining the present spatial extent include prey availability, migration patterns and land use change. Recently, the dramatic decline in insects has drawn attention to insectivorous predators [BirdLife Schweiz 2018; Sumser et al. 2017]. Many insectivorous ground breeders are threatened not only by this deficit of nourishment but also by the loss of habitat, especially in the intensified agricultural landscape [Sánchez-Bayo and Wyckhuys 2019; Schrauth and Wink 2018]. For example, breeding success and survival of the northern wheatear (*Oenanthe oenanthe*) depends more on the habitat quality than on the individual's quality [Arlt and Part 2007; Currie et al. 2000]. Therefore, habitat suitability analysis for northern wheatears has gained interest. Recent studies predicted severe habitat loss [Chamberlain et al. 2013; van Oosten et al. 2014], whereas a lack of suitable habitat was not (yet) an issue in the 1990s [Tye 1992].

The male northern wheatears claim a suitable breeding ground after the exhaustive journey from their wintering quarters [Arizaga and Bairlein 2011; Currie et al. 2000; Maggini and Bairlein 2012]. Almost all northern wheatears overwinter in sub-Saharan Africa, which delineates them as long-distance migrants [Cramp et al. 1988; Maumary et al. 2007; Öberg 2014; Schweizerische Vogelwarte Sempach n.d.]. Their preferred nest sites in Europe feature open, heterogeneous fields with short vegetation [Arlt and Part 2007; Olliver et al. 1999; Tye 1992; van Oosten et al. 2014]. The habitat choice of this European breeding population differs slightly from that of the population in the European Alps. The Alpine northern wheatears choose mostly similarly (1) open, (2) heterogeneous and (3) short-vegetated sites, which favours their hunting strategy [Conder 1989; Hussell et al. 2014; Tye 1992; van Oosten et al. 2014]. These nest sites are located on (4) SW-exposed slopes of (5) a less than 20° incline, with the steepest territory being on a slope of a more than 40° incline [Wartmann 1985]. Nests were built usually (6) around or above the treeline, with a maximal density between 2200 and 2500m a.s.l. [Glutz von Blotzheim 1962; Glutz von Blotzheim and Bauer 1988; Knaus et al. 2018]. The chosen sites are (7) snow-free before 31th May and contain (8) rocks, fields of debris, cairns or other convex features [Glutz von Blotzheim and Bauer 1988; Mattes et al. 2005; Olliver et al. 1999; Pollheimer and Gstir 2013; Schmid et al. 1998; Wartmann 1985]. The local distribution of nests was described as (9) small clusters [Arlt and Part 2007] or with a minimal recorded distance between two nests of barely 100m [Glutz von Blotzheim 1962]. The found average number of breeding pairs ranged between 2.5/km^2 and 5/10ha [Catzeflis 1976; Glutz von Blotzheim 1962]. No literature values of the distance from a northern wheatear nest to the nearest path exist to my knowledge.

While occasional breeding records exist about the Swiss lowlands until the 1970s, the nest sites of northern wheatears are nowadays restricted to the Alpine areas [Knaus et al. 2018; Maumary et al. 2007]. This Alpine population is currently stable or even slightly increasing [Knaus et al. 2018]. Concurrently, the climate is changing, the Alpine farming developing and the suitable habitats are expected to decline [Chamberlain et al. 2013; Sánchez-Bayo and Wyckhuys 2019; Schrauth and Wink 2018; van Oosten et al. 2014]. Yet, little research on the characteristics of suitable nest sites has been undertaken recently. Therefore, I aim to determine which environmental factors affect the nest site choices of northern wheatears in the Piora valley, Canton of Ticino, Switzerland. As Alpine environments alter frequently because of avalanches and soil erosion processes, I also aspire to determine to what extent the choice is influenced by the large between-year variation of the analysed environmental factors.

The more complete understanding of the patchy Alpine distribution demands a certain consistency of these factors [Meier 2018]. The previously studied snow coverage, slope, elevation, aspect, land cover

and clusters might also describe the population in Val Piora [Arlt and Part 2007; Pollheimer and Gstir 2013; Wartmann 1985]. The literature described the suitable breeding grounds further. However, the sampled data prohibited the analysis of a particular factor, such as the vegetation density of the nest site. A similar variable, the land cover, was not available for the data points in the Val Piora. Therefore, I aimed to substitute or approximate this factor by an effortlessly computable variable, such as the normalized difference vegetation index (NDVI). I investigated the potentials and limitations of a substitution of NDVI and the sampled land cover. Additionally, I extend the previous investigation by the cumulated distance to the three nearest nests, the distance to the nearest path, the NDVI and the roughness of the terrain. In the Val Piora, I expect northern wheatears to nest on lower altitudes which are assumedly earlier snow-free. I assume that northern wheatears prefer sites with little variation in slope and that they almost completely avoid perpendicular sites. Furthermore, I suspect northern wheatears to favour hillsides which are almost exclusively S-exposed. I further suppose that northern wheatears avoid wide concave structures such as depressions because snow persists there. On the other hand, I expect northern wheatears to nest rather on narrow convex structures such as a cairn as I think these compositions are earlier snow-free. Additionally, I assume that the nest sites display some heterogeneity, which could be indicated either by alternating convex and concave structures or by varying land covers. Latter implies considerable NDVI variation within the nest site. Moreover, I suppose northern wheatears prefer locations with slightly positive NDVI values which indicate sparse or no vegetation to favour their hunting strategy. I expect northern wheatears to not nest next to each other because of the competition for prey while I further predict them to not scatter evenly over the study area. Furthermore, I expect the northern wheatears to either avoid disturbances and keep a distance to the nearest path or to favour a certain proximity to paths as they were built on less steep terrain [Bötsch, Tablado, and Jenni 2017; Bötsch, Tablado, Scherl, et al. 2018; Finney et al. 2005]. Generally, I suppose northern wheatears prefer nest sites which are snow-free for the first time early in season and whose last day of snow is only days later. I expect the first snow-free day and the last snow-covered day to approximate the biological availability of a nest site.

The recent changes in Alpine snow dynamics has generated wide interest in snow-related processes [Bavay et al. 2013; Bormann et al. 2018; Scherrer et al. 2004]. According to Wartmann 1985, northern wheatears preferably nested on sites snow-free by the end of May. If this schedule could be confirmed, it would provide clear evidence that the nest site selection of the northern wheatears in the Swiss Alps depends on the snow dynamics. Thus, it is worth investigating how the snow coverage changes at various nest sites during the year. I expect sites to be earlier snow-free if they lie on southward-exposed convex structures at lower altitudes. Therefore, I wrote an algorithm in R to semi-automatically extract the snow coverage for various nest sites from ordinary RGB digital camera pictures with the help of a digital elevation model.

The method section starts with background information on the study object: the northern wheatear. Then, I introduce the study area and the data used for modelling. The compilation of all needed variables was time-consuming, especially those describing the snow coverage (SC) which I thus describe extensively. The method chapter ends with the fitted Bayesian models. Subsequently, I present the obtained results, first those of the SC algorithm and finally those of the models. Naturally, some difficulties arose and consequently, decisions had to be taken along the path. These obstacles and their overcoming are handled in the discussion section, which continues with the implications of the results of the SC algorithm and models. Furthermore, the results are compared to previous research. Lastly, the discussion finishes with an outlook on possible improvements and future steps.

2 METHODS

2.1 The study object - northern wheatear (*Oenanthe oenanthe*)

2.1.1 Distribution

The breeding areas of the northern wheatear (*Oenanthe oenanthe*) span almost circum-Arctic with gaps only in the centres of Russia, Greenland and North America [Birdlife International 2018; Cramp et al. 1988; Heath et al. 2000]. Four different subspecies live in this vast geographic range with the nominal form occuring in mainland Europe, Siberia, British Islands and Alaska [Glutz von Blotzheim and Bauer 1988; Maumary et al. 2007]. While occasional breeding records exist in the Swiss lowlands from before the 1970s, the nest sites of northern wheatears are nowadays restricted to the Alpine areas [Knaus et al. 2018; Schmid et al. 1998; Winkler 1999] (see Figure 1). The highest-situated nest was discovered at 2880m a.s.l. near Zermatt [Glutz von Blotzheim and Bauer 1988]. This Alpine population is currently stable, despite substantial fluctuations; it may even be slightly increasing [Knaus et al. 2018].
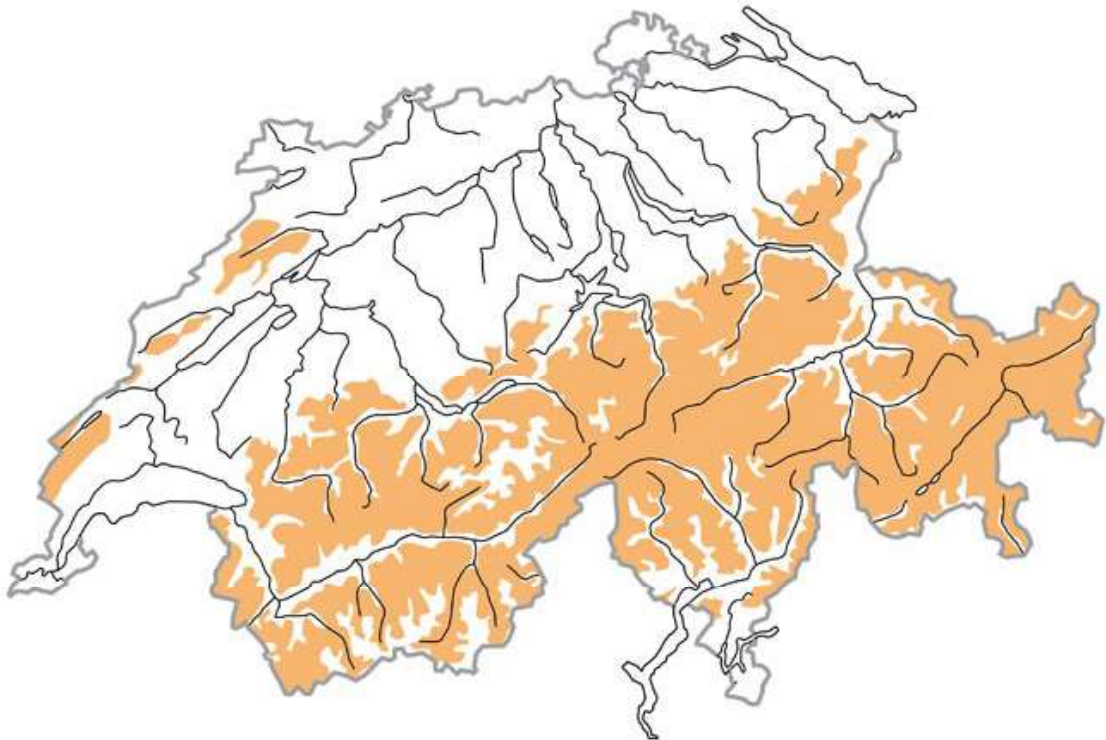
Figure 1: In Switzerland, northern wheatears remain in mountainous areas during breeding season and desert the country completely in winter [Schweizerische Vogelwarte Sempach n.d.].

2.1.2 Migration strategy

Most northern wheatears desert their Alpine breeding grounds between the end of August and mid-September [Glutz von Blotzheim 1962]. Similarly to almost all other populations, Alpine northern wheatears are long-distance migrants and overwinter in sub-Saharan Africa [Cramp et al. 1988; Glutz von Blotzheim and Bauer 1988; Öberg 2014; Schweizerische Vogelwarte Sempach n.d.]. Eventually, they head back towards their breeding grounds. Those returning to Switzerland arrive generally between

the end of April and start of May at their breeding grounds [Glutz von Blotzheim 1962; Glutz von Blotzheim and Bauer 1988].

2.1.3 Brood biology

The northern wheatears in the Alps rear their young for most of their time at the breeding grounds. The breeding pair builds their nest in cavities of preferably stone walls usually not higher than one metre above ground towards mid-May [Corti and Melcher 1958; Glutz von Blotzheim 1962; Glutz von Blotzheim and Bauer 1988; Wartmann 1985]. A nest protects the offspring from weather and predators [Meier 2018]. Then, the female lays a clutch of normally five to six eggs around the end of May or start of June [Corti and Melcher 1958; Glutz von Blotzheim 1962]. Incidentally, the recent Swedish breeding seasons have commenced earlier because of warmer springs [Pärt et al. 2017]. If the first brood fails or if it succeeds early in season, the female lays a second clutch in a fresh nest between the end of June and start of July [Glutz von Blotzheim 1962]. The young of these late broods fledge at the start of August, though records exist of later fledglings [Corti and Melcher 1958; Glutz von Blotzheim 1962].
Northern wheatears show rather site than mate constancy over consecutive seasons [Brooke 1979; Currie et al. 2000]. Indeed, the only recorded reformation of a pair was on a site where this pair bred together before [Brooke 1979]. However, suitable territories are reclaimed if they are available [Brooke 1979; Currie et al. 2000]. This site constancy is not observed in cases of a previously failed brood as the unsuccessful breeder will try to secure a more productive territory in such a case [Arlt and Pärt 2008b]. Typically, the most suitable territories are occupied first by different and usually older individuals [Arlt and Pärt 2008b; Brooke 1979; Currie et al. 2000]. Experienced breeders select their breeding ground in "a two-step process, made partly in the post-breeding period and partly at the time of territory establishment in the subsequent year" [Arlt and Pärt 2008a]. Northern wheatears are described as territorial birds with defended territories in winter quarters, migration stopover sites and breeding areas [Maumary et al. 2007; Tye 1992]. However, they are observed to nest relatively close together [Corti and Melcher 1958] or in small clusters [Arlt and Part 2007]. They appear to defend mostly the nest site and the nesting partner whereas they are less territorial during foraging [Meier 2018].
The habitat choice of the European breeding population differs slightly from that of the population in the European Alps. Preferred nest sites in Europe are mainly in the lowland and feature open, heterogeneous fields with short vegetation [Arlt and Part 2007; Glutz von Blotzheim and Bauer 1988; Olliver et al. 1999; Tye 1992; van Oosten et al. 2014]. On the other hand, Alpine northern wheatears choose such open, heterogeneous and short vegetated sites on SE to SW exposed slopes of less than 30°. They nest preferably around or above the treeline, on locations which are snow-free before 31th May and contain rocks, fields of debris or cairns [Glutz von Blotzheim 1962; Mattes et al. 2005; Maumary et al. 2007; Pollheimer and Gstir 2013; Schmid et al. 1998; Wartmann 1985].

2.1.4 Foraging strategy

The dominant foraging strategy of the northern wheatear is described as "dash-and-jab", which is adapted to short, sparse vegetation [Conder 1989; Hussell et al. 2014; Tye 1992; van Oosten et al. 2014]. Another approach is "sallying" where the prey is spotted from a perch, from which the chase starts and to which the hunter returns eventually [Glutz von Blotzheim and Bauer 1988; Hussell et al. 2014; Pärt 2001]. The thus collected diet mostly consists of insects, spiders, worms and small snails, which is supplemented with berries in fall [Glutz von Blotzheim 1962; Glutz von Blotzheim and Bauer 1988; Hågvar et al. 2009; Tye 1992; van Oosten et al. 2014; Wartmann 1985].

2.2 The study area - Val Piora, Canton of Ticino, Switzerland
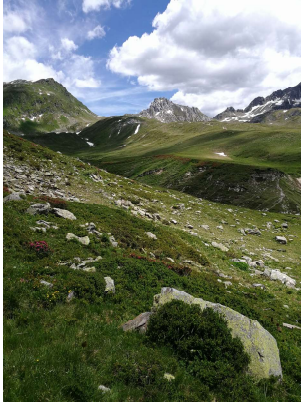
Figure 2: The view in direction of Pizzo del Sole on 29th June 2018.

The study area covers ~10 km^2 of Val Piora, Canton of Ticino, Switzerland. This picturesque Alpine valley starts at 1850m a.s.l., spans around well-frequented hiking trails amongst the four mountain lakes Lago Ritom, Lago Cadagno, Lago di Tom and Lago di Dentro. The latter three lakes and the river Murinascia Grande with its tributaries discharge into the storage reservoir Lago Ritom [Swisstopo 2019]. The glacial origin of this landscape is still noticeable [Bachofen et al. 2016]. More recent structures are screes, which are often found at the bottom of cliffs. A debris field gathered to cairns lies on the northern slope between Lago Cadagno and Lago di Dentro. Such S-exposed inclines are often pastures, whereas northern faced slopes are dominated by "blueberries, rhododendron-vaccinion and juniperus" [Bachofen et al. 2016]. These opposing hillsides are separated by the Murinascia Grande, which flows in the lower part through deep gorges.
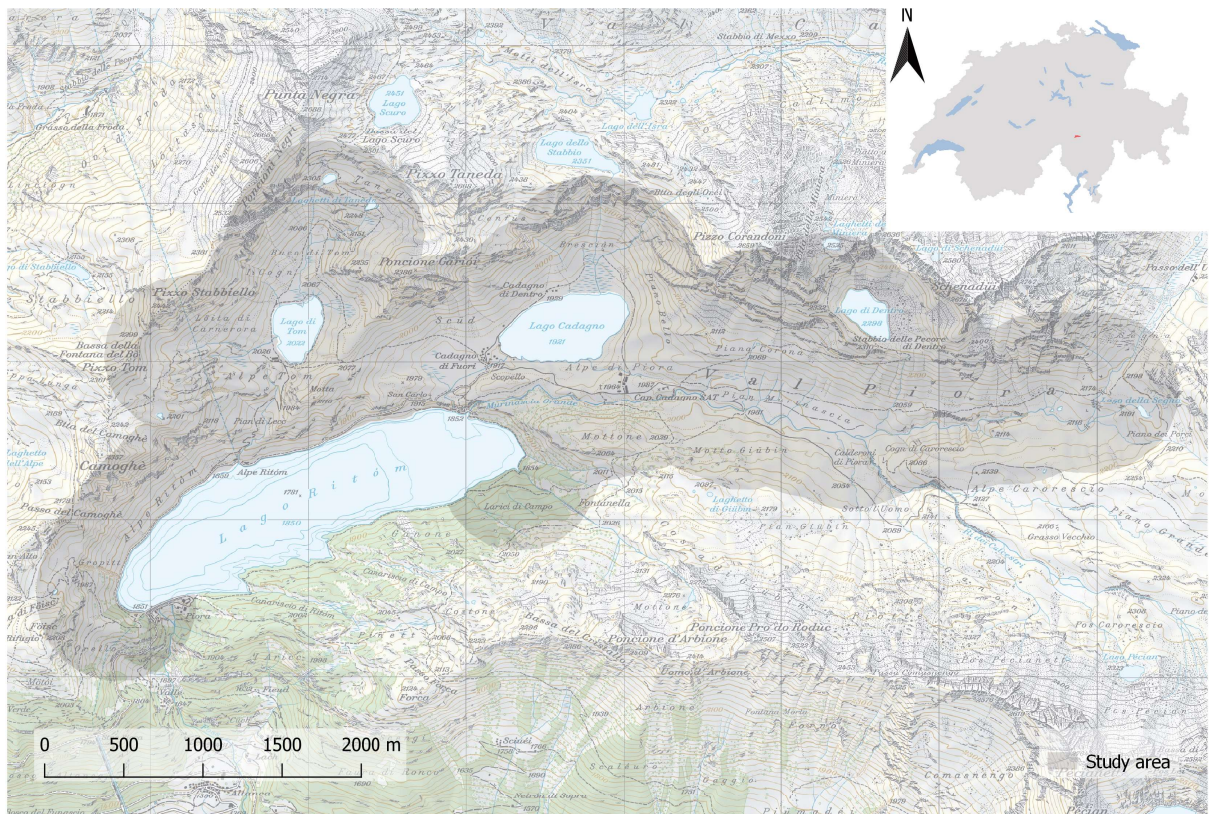
Figure 3: The study area in the Val Piora is located in the northern part of the Canton of Ticino, Switzerland (red area on the small map). It stretches around hiking trails and lakes (semi-transparent area on the main map).

2.3 The data

The Swiss Ornithological Institute (SOI) has collected data concerning nests and nest sites of northern wheatears in Val Piora since 2011. A nest and a nest site both are important parameters to the breeding northern wheatear. The coordinates of a nest refer to the cavity where the eggs are laid. This point in space locates the centre of the nest site. The nest site refers here to the surroundings of the nest. I assumed that the nest site was chosen as breeding ground because of the characteristics of the environmental factor in its extent. To be precise, the nest site is here empirically defined as the area within a radius of 25m around the nest.

The gathered variables describe four different kinds of information of the nest site: position, breeding success, surrounding features and data on snowmelt. The position locates the nest on the Swiss Grid (LV03) and allowed the location-specific computation of environmental factors. The breeding success was defined as the presence of a nest (presence point). The analysed surrounding features in this thesis were: elevation, slope, aspect, roughness, normalized difference vegetation index (NDVI), cumulated distance to the three nearest nests and distance to the nearest path. The raw data collected in the field yet not analysed environmental factors were: land cover, cavity abundance, perch abundance, most frequent perch height, maximal perch height and cairn abundance. While a team of SOI sampled the land cover between May and July, see 6.2.2, I qualitatively recorded the cavity abundance, perch abundance, most frequent perch height, maximal perch height and cairn abundance, see 6.2.9, 6.2.10 and 6.2.11. The snow coverage variables describe different milestones of the availability of the nest site for the breeding pair. The computed snow coverage variables were first snow-free day, last snow-covered day and number of temporary snow-free days. The first snow-free day approximates the earliest possible date when birds got access to cavities on the ground. The last snow-covered day describes the latest snow-induced constraint. The number of temporary snow-free days describes the availability of the site during the snowmelt. Their computation required the footage of a stationary digital camera installed by a team of SOI taking a picture every hour. The used cameras was the GardenWatchCam by brinno and the HYPERFIRE by RECONYX. The here presented snow coverage algorithm, see 2.4, was restricted to the photos of the GardenWatchCam. The GardenWatchCam and the HYPERFIRE displayed different spectral behaviours. Thus, two slightly different methods would have been required, which was unrealistic in the given time. Therefore, I focused on the camera which recorded the major part of the data and which still is in use.

A total of 334 nests were sampled, the land cover was assessed on 232 nest sites. I recorded the data concerning perches, cavities and cairns of 275 nest sites in 2018, even for nests observed between 2012 and 2017. The snow coverage was calculated for 37 nests. In general, the data was collected in the year of the observed nest. Note that not all factors were available for every nest site or every year. The variables describing land cover, perches, cavities and cairns were only available for presence points but not for absence points.

The above described data was collected to model the likelihood of a nest for a specific location. The location, i.e. nest site, was described by the computed environmental factors and snow coverage. A location was either occupied by a nest (presence point) or not occupied (absence point). Thus a binomial model was required. I expected varying effects of the sampled environmental factors, see above. As binomial nest site suitability models require presence and absence data [F. Korner-Nievergelt and P. Korner-Nievergelt 2019], I created 241 pseudo-absence data points to match the original number of cleaned presence data, see 2.3.1.

2.3.1 Creation of pseudo-absence points

The pseudo-absence points created here mark sites where the field assistants of the SOI have never recorded a nest. The pseudo-absence points (hereafter absence points) enable the fitting of the later used generalized linear models (GLM) [F. Korner-Nievergelt and P. Korner-Nievergelt 2019]. The suggested number of absence points has been ten times the presence points (i.e. nests) [Chefaoui and Lobo 2008; Kanagaraj et al. 2013]. However, given the limited size of the study area and the following limitations in placement, this number of absence points would have overcrowded the available area. Therefore, I decided to create as many absence points as presence points.

The absence points spread generally randomly across the study area. The study area spans within 500m of the patrolled paths, see the semi-transparent grey polygon in Figure 4. To avoid trivial results, I excluded impossible nest locations like waterbodies[1] and forests[2], see the light blue and light green polygons in Figure 4 [Meier et al. 2018]. To prevent replications, I further excluded locations which were nearer than 50m to a nest [Meier et al. 2018]. If the absence point was chosen within a radius of 50m around the nest, the site around this absence point would overlap with the nest site. Then, I chose 241 absence points in the remaining study area using *Random points inside polygon (fixed)* in QGIS with a minimal distance of 50m between two points. Finally, I assigned the years 2013 to 2018 to the absence points with the frequency distribution found in the sampled nest sites.
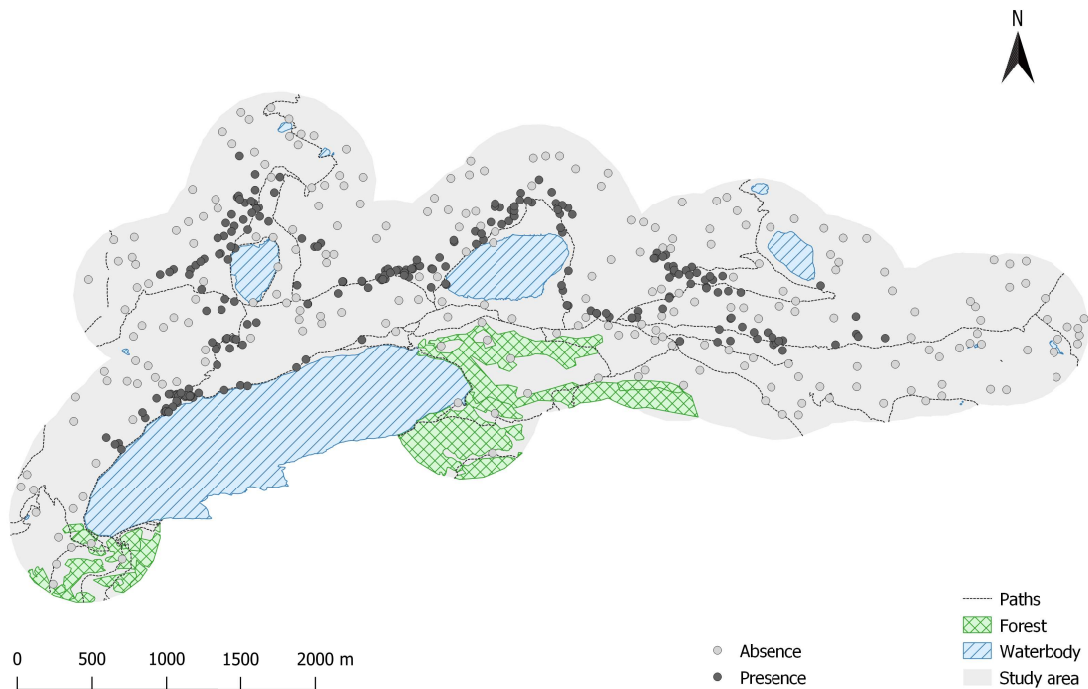

Figure 4: The dark grey presence points mark nest sites and the absence points are light grey. They spread across the study area, which includes four lakes, forests and hiking paths.

---

[1] The layer Waterbody origins from swissTLM3D - Gewaesser by swisstopo.

[2] The layer Forest origins from Vector25 - Primaerflaechen by swisstopo.

2.3.2 Moving window technique on raster data

Each of the presence and absence points was described by a set of environmental factors. These factors were either vector or raster data. Vector data describe a feature as either point, line or polygon [QGIS n.d.], e.g. coordinates of a nest. Raster data generally describe area-covering cell-based variables [QGIS n.d.], e.g. elevation, slope, aspect, roughness, NDVI and distance to the nearest path.

The considered environmental factors were to describe the nest site rather than the location of the nest. Thus, multiple raster cells were required to compute the overall characteristics of the displayed factor for one nest site. Remember, the nest site is defined as the area of a circle with radius 25m. The number of cells to analyse one nest site depends on the spatial resolution of the raster. I calculated a measure of central tendency and a measure of dispersion from the cells of the nest site [Meier et al. 2018]. The measure of central tendency, here median, describes the overall appearance of the nest site rather than the punctual value of the nest [Meier et al. 2018]. Latter might coincidentally correspond to a cell with an extreme value. This process leads to a smoothed raster. The measure of dispersion, here 90% percentile distance, describes the variability within the nest site [Meier et al. 2018]. This process leads to a raster with enhanced edges. The 90% percentile distance equals the 95% percentile minus the 5% percentile.

I calculated the median and the 90% percentile distance in R using the function *raster::focal*. This application requires among others a square matrix with an odd number of rows defining the neighbourhood to analyse. Whereas I need a circular matrix (moving window) which defines the neighbourhood as the cells within a 25m radius around the central cell to analyse the study area despite varying spatial resolutions[3] of the input rasters. Therefore, I created a function to produce a matrix like Figure 5.

```
.  .  .  .  .  .  .  .  .  .  .  1  .  .  .  .  .  .  .  .  .
.  .  .  .  .  .  .  1  1  1  1  1  1  1  1  1  .  .  .  .  .
.  .  .  .  .  1  1  1  1  1  1  1  1  1  1  1  1  1  .  .  .
.  .  .  1  1  1  1  1  1  1  1  1  1  1  1  1  1  1  1  .  .  .
.  .  1  1  1  1  1  1  1  1  1  1  1  1  1  1  1  1  1  1  .  .
.  .  1  1  1  1  1  1  1  1  1  1  1  1  1  1  1  1  1  1  .  .
.  1  1  1  1  1  1  1  1  1  1  1  1  1  1  1  1  1  1  1  .
.  1  1  1  1  1  1  1  1  1  1  1  1  1  1  1  1  1  1  1  .
.  1  1  1  1  1  1  1  1  1  1  1  1  1  1  1  1  1  1  1  .
.  1  1  1  1  1  1  1  1  1  1  1  1  1  1  1  1  1  1  1  .
1  1  1  1  1  1  1  1  1  1  C  1  1  1  1  1  1  1  1  1  1
.  1  1  1  1  1  1  1  1  1  1  1  1  1  1  1  1  1  1  1  .
.  1  1  1  1  1  1  1  1  1  1  1  1  1  1  1  1  1  1  1  .
.  1  1  1  1  1  1  1  1  1  1  1  1  1  1  1  1  1  1  1  .
.  1  1  1  1  1  1  1  1  1  1  1  1  1  1  1  1  1  1  1  .
.  .  1  1  1  1  1  1  1  1  1  1  1  1  1  1  1  1  1  .  .
.  .  1  1  1  1  1  1  1  1  1  1  1  1  1  1  1  1  1  .  .
.  .  .  1  1  1  1  1  1  1  1  1  1  1  1  1  1  1  .  .  .
.  .  .  .  1  1  1  1  1  1  1  1  1  1  1  1  1  .  .  .  .
.  .  .  .  .  .  1  1  1  1  1  1  1  1  1  .  .  .  .  .  .
.  .  .  .  .  .  .  .  .  .  .  1  .  .  .  .  .  .  .  .  .
```

Figure 5: A moving window of 21 x 21 cells. A dots refers to NA values and C marks the central cell.

There are four particularities worth noting in the moving window in Figure 5: First, the NA values allow to exclude those cells from the analysis. Therefore, the square matrix becomes circular. Second, the central cell C usually takes the value 1. The special case of C = -1 applies only for the calculation of the 90% percentile distance of the aspect as described in the succeeding paragraph. Third, the ratio of the radius [m] to the cell size [m] gives the radius in cells. Let c1 be any cell in Figure 5 except the

---

[3] The cell size or spatial resolution of the rasters digital elevation model (DEM), slope, aspect and roughness is 2m whereas the cell size of the NDVI raster is 0.25m.

central cell. A cell c1 is assigned 1 if the linear distance between c1 and C is smaller than or equal to the radius (in cells) and NA otherwise. Thus, I considered only those cells for the computation of the median or the 90% percentile distance whose actual distance to the cell containing the nest was 25m or less. Fourth, the number of cells per row changes according to the spatial resolution of the raster. The shape of a circle given by C and the radius is approximated more accurately with a higher spatial resolution. The drawback of a more circular moving window is the longer calculation time because of the increased number of cells to analyse.

The 90% percentile distance of the aspect required a slightly different approach. Let us consider two nest sites. The aspects of a N-exposed nest site range between >350° and <10°. The aspects of a south-faced nest site range between >150° and <210°. The maximal range in aspect of the south-faced site is easily calculated as 210° - 150°. The same calculation fails for the north-faced nest site because of the periodical nature of the aspect; North equals both 0° and 360°. I decided to offset C to always correspond to 180°, meaning to add x°. By adding the same x° to all other values of the particular moving window, I maintained their distribution in relation to C [Messmer 2019]. Consequently, the central cell needs to be addressed and its value extracted. Otherwise x remained unquantified, which prohibited the correct offset computation. To address the central cell during the computation, I assigned -1 to C, which makes the central cell negative (e.g. C = -30° instead of 30°), while all other values are either positive or NA.

$$offset = (-1) * C + 180 \tag{1}$$

The offset corresponds to this negative value + 180°, see formula 1 (e.g. offset = -30°+ 180° = 150°). Then, I changed the central cell back to its true value (e.g. C = 30°).

$$c1_{\text{new}} = (c1_{\text{old}} + offset)\%\%360 \tag{2}$$

Finally, I added the offset to all cells of the matrix and brought the values back into the original range, see formula (2) (e.g. c1 = (270° + 150°) %% 360 = 420° %% 360 = 60°). This preparation step allows now to apply the 90% percentile distance function on the aspect raster.

Variable	Unit	Value range and interpretation
Elevation	m a.s.l.	The elevation in metre above sea level.
Slope	degree	The slope reports the steepness of the ground. 0° signifies horizontal terrain and 90° signifies vertical terrain.
Aspect	degree	The aspect characterizes the orientation of the terrain. A hillside of aspect 0° (or 360°) is N-exposed, 90° is E-exposed, 180° is S-exposed and 270° is W-exposed. The aspect is only calculable for slopes >0°.
Roughness		The roughness indicates here the topography with values ranging from -1 to +1. Zero indicates smooth terrain, negative values ridges and positive values valleys. Large absolute values indicate narrow features while small absolute values indicate wide structures.
Normalized difference vegetation index (NDVI)		The NDVI discriminates three land cover classes: (1) a negative index suggests water, clouds or snow, (2) an index around zero suggests bare soil or rock and (3) a positive index suggests vegetation [Lillesand et al. 2008]. The closer the NDVI to one is, the greener or denser the vegetation is.
Cumulated distance to the three nearest nests	metre	A small cumulated distance to the three nearest nests values hint at a clustering of nests. Negative values are not possible.
Distance to the nearest path	metre	The distance to the nearest path corresponds to the minimal walking distance to the next path. Negative values are not possible.
First day snow-free	day	An area is considered snow-free if <10% of the surface is covered in snow. The first snow-free day equals the first day where this condition is met. The first of January of any year equals day one.
Last day snow-covered	day	The last snow-covered day equals the latest day where ≥10% of the surface of the nest site is covered in snow. The first of January of any year equals day one.

Table 1: A summary of all in a statistical model used variables, their unit of measurement and the interpretation of their value range. The variables slope, DEM, aspect, clusters/cumulated distance to the three nearest nests, distance to the nearest path, NDVI and roughness are deducted from remotely sensed raster data. Whereas the snow coverage (first snow-free day and last snow-covered day) is extracted from the combination of a digital elevation model and ordinary photographs. All of these variables were analysed using GLMs in R.

Please note: For each nest site, two measures exist for the variables elevation, slope, aspect, roughness and NDVI, see 2.3.2. The median value was computed as a measure of central tendency and the 90% percentile distance (95% percentile minus 5% percentile) as a measure of dispersion. Consequently, the variables are either called median of *variable* or 90% percentile distance of *variable*. These variable names especially affects the analysed data from 2.3.3 to 2.3.7 which was also included in the models in 2.5.1 and 2.5.2. The original raster layers are visible in 6.2.

Variable	Unit	Value range and interpretation
Number of temporary snow-free days	day	The number of temporary snow-free days counts the number of snow-free days between the first snow-free day and the last snow-covered day, see Table 1.
Land cover	percent	Land cover consists of the categories (1) vegetation <5cm, (2) vegetation 5 - 15cm, (3) vegetation >15cm, (4) shrubs, (5) bare ground and (6) rocks. 20% shrubs implies that 20% of the surface within a radius of 25m was covered in shrubs. The total land cover per nest site is 100%. The sampling date varies strongly within the season.
Cavity abundance		The cavity abundance expresses qualitatively the number of cavities available at a specific nest site in June 2018 using the categories (1) none, (2) few, (3) some and (4) many. A cavity is any kind of niche that protects from wind, rain and predators.
Perch abundance		The perch abundance expresses qualitatively the number of perches available at a specific nest site in June 2018 using the categories (1) none, (2) few, (3) some and (4) many. A perch overlooks the surrounding.
Most frequent perch height	metre	The most frequent perch height assesses qualitatively the most often encountered height of a perch at a specific nest site in June 2018 using the categories (1) <0.5m, (2) 0.5 - 1.0m and (3) >1m.
Maximal perch height	metre	The maximal perch height assesses qualitatively the highest encountered perch at a specific nest site in June 2018 using the categories (1) <0.5m, (2) 0.5 - 1.0m and (3) >1m.
Cairn abundance		The cairns abundance expresses qualitatively the number of cairns available at a specific nest site in June 2018 using the categories (1) none, (2) one or two and (3) three or more. A cairn is a pile of stones built from collected rocks in the surroundings.

Table 2: A summary of variables I collected in the field, but which were excluded from further statistical analysis, their unit of measurement and the interpretation of their value range.

2.3.3 Digital elevation model (DEM)

– Source: swissALTI^{3D}
– Sampling date: 2012/08/19 [Swisstopo n.d.(a)]
– Spatial resolution: 2m
– Maximal error per data source [Swisstopo n.d.(a)]:
    – up to 2000m a.s.l., ± 0.5m in all dimensions
    – above 2000m a.s.l., ± 3m
– Maximal error[4]: ± 3m
– Computation method: none
– Remarks: When the northern wheatear arrived in the Val Piora in spring, the Lago Ritom was partially empty. Some northern wheatears chose the exposed rocky beaches of the dam as breeding grounds. On the other hand, the Lago Ritom was completely filled when the digital elevation model (DEM) was recorded. Thus, the nests on its shores are perceived to be in or under water. The true elevation of the nest would have been lower than the lake level. The flat surface of the lake not only created wrong values for slope and roughness, it also prohibited the computation of the aspect. Thus, I decided to exclude the nest sites on the shore of the Lago Ritom from further analysis.
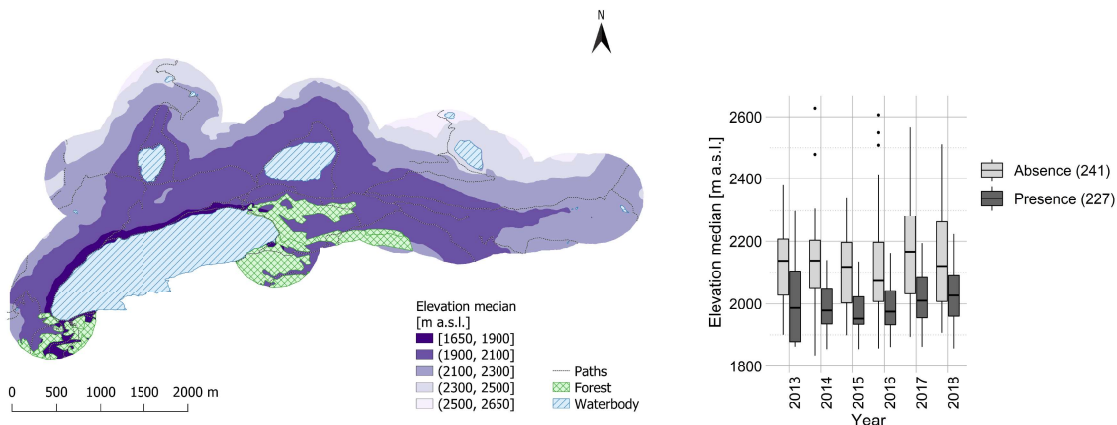
(a) The median of the elevation in the study area ranges between 1650m a.s.l. and 2650m a.s.l. The most elevated areas are found in the north. The low plain around Lago Cadagno and Lago di Tom dominates the study area.

(b) The presence points are located often lower, e.g. at the north shores of the three biggest lake. This accordance with my expectation opposes the literature values.

Figure 6: However low situated, the entire Val Piora is almost free of trees because of human activities [Meier 2018]. This might indicate that the northern wheatear prioritise sparse vegetation over altitude.

---

[4] Maximal error of the created raster. Consider the additional spatial inaccuracy of the recorded GPS signal, see 6.2.1

2.3.4 Slope

– Source: swissALTI^{3D} for elevation
– Sampling date: 2012/08/19 [Swisstopo n.d.(a)]
– Spatial resolution: 2m
– Maximal error per data source [Swisstopo n.d.(a)]:
    – up to 2000m a.s.l., ± 0.5m in all dimensions
    – above 2000m a.s.l., ± 3m
– Maximal error[4]: ± 3m
– Computation method: I calculated the slope with the function *Slope, aspect, curvature* in QGIS and the specification *9 parameter 2nd order polynom (Zevenbergen & Thorne 1987)* on above DEM.
– Remarks: none
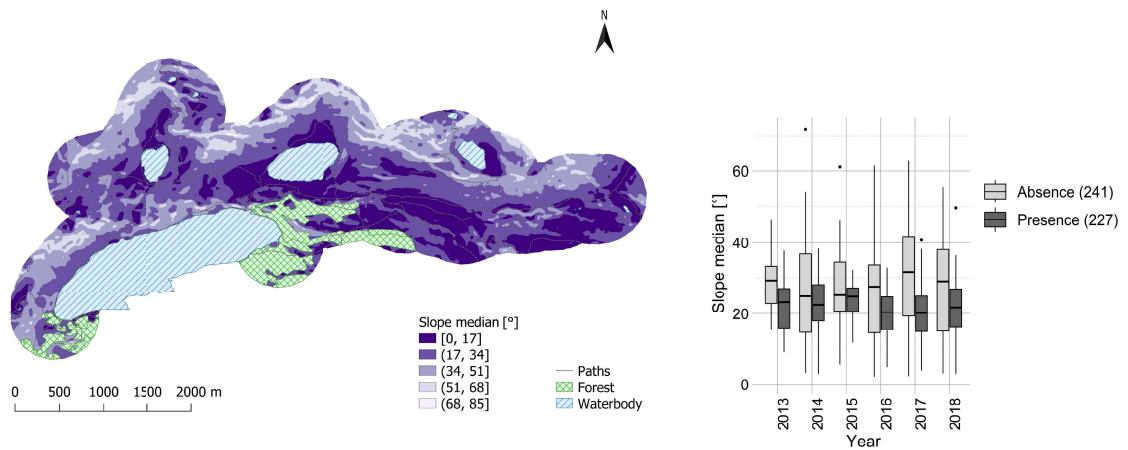
(a) The flat valley bottom in the north-east contrasts the almost vertical cliffs in the north.

(b) The presence and absence points overlap in the flatter terrain. Hardly any nests are found on steep slopes.

Figure 7: The median of the slope could theoretically range between 0° and 90°, which the values in the study area almost reach. Yet, most nest sites are no steeper than 30° as the literature suggested. On the other hand, horizontal plains seem to be less chosen despite their availability which opposes my assumption.
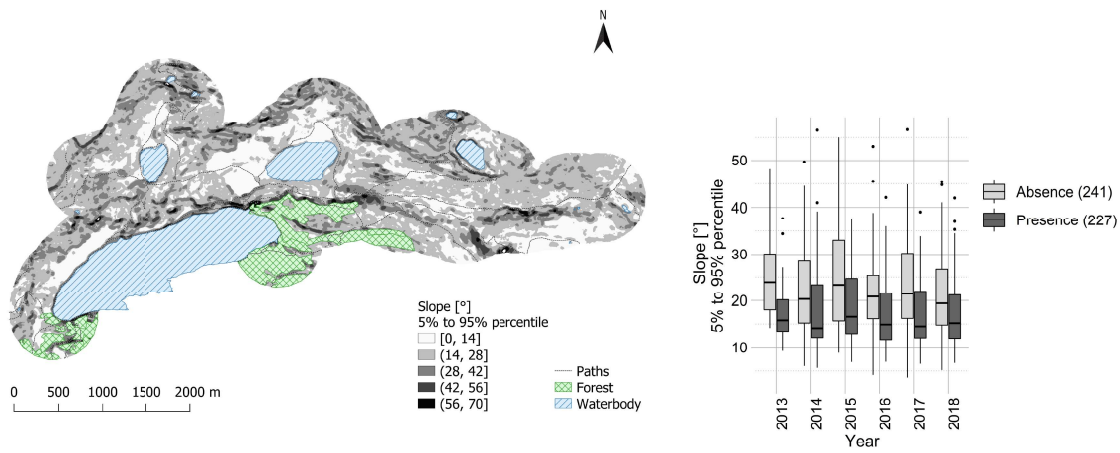
(a) The most uniform slopes span around the north-west shore of Lago Ritom, Lago Cadagno and Lago di Tom. Generally, the Val Piora ranges less than 28°.

(b) The presence points are generally in more uniform areas than the absence points. Nonetheless, a 90% percentile distance of 20° is immense if the median value itself is 30°.

Figure 8: The 90% percentile distance of the slope could theoretically range between 0° and 90°, although the range of the study area remains <70°.
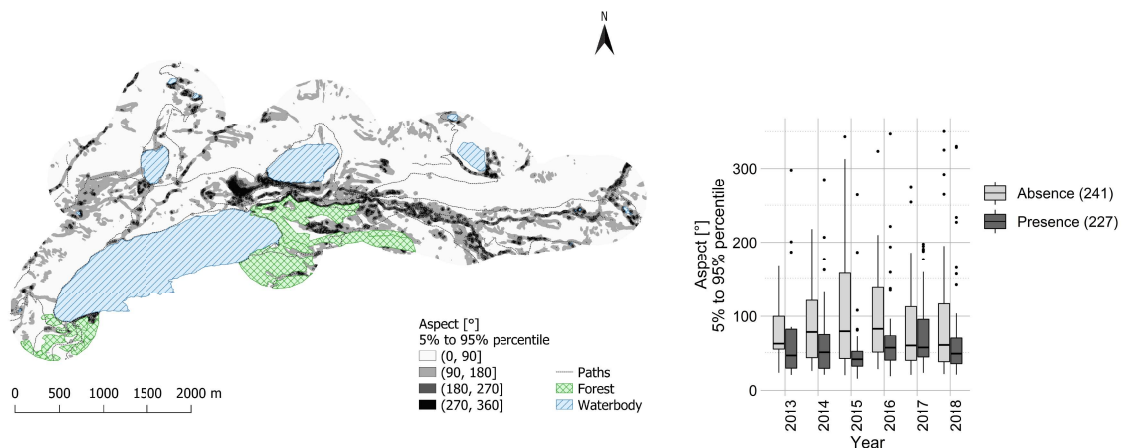
2.3.5 Aspect

– Source: swissALTI^{3D} for elevation
– Sampling date: 2012/08/19 [Swisstopo n.d.(a)]
– Spatial resolution: 2m
– Maximal error per data source [Swisstopo n.d.(a)]:
    – up to 2000m a.s.l., $\pm$0.5m in all dimensions
    – above 2000m a.s.l., $\pm$3m
– Maximal error[4]: $\pm$3m
– Computation method: The aspect is computed in the same tool as slope but it is undefined for absolutely flat areas (slope = 0°).
– Remarks: none
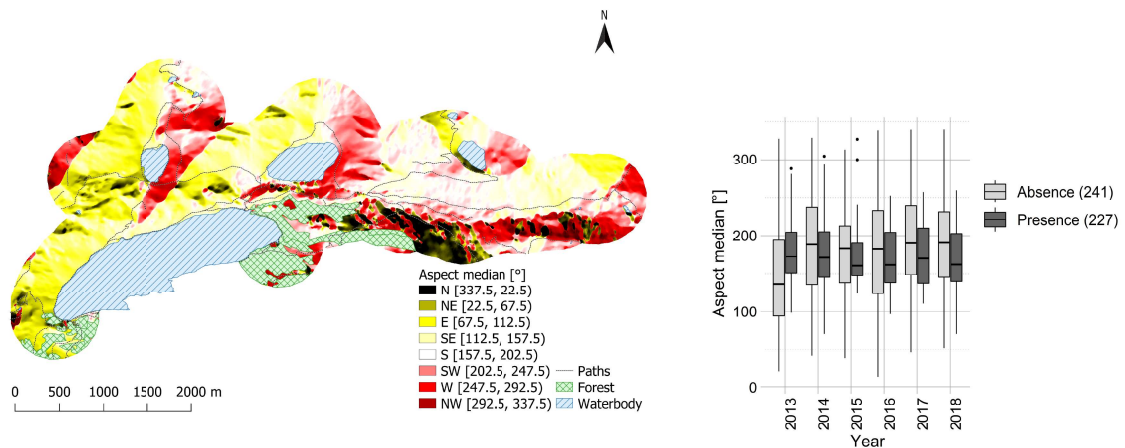
(a) While the western part of the study area divides into E-exposed and west-faced, the eastern part divides into south-faced, north-faced and west-faced.

(b) The presence points vary less in the orientation of the terrain and range mostly between SE and SW, meeting my expectations.

Figure 9: The median of the aspect could theoretically range between 0° and 360°, which it does in the study area.

(a) The aspect of the study area is mostly uniformly exposed. The dark areas range up to the highest possible values and correspond to the ridges and river bed.

(b) The presence points are generally on more uniform inclines and vary less.

Figure 10: The 90% percentile distance of the aspect could theoretically range between 0° and 360°. The found variation in aspect was according to my own expectation but also corresponds to the dominant value of the study area. Therefore, the explanatory power of this variable may be limited.

2.3.6 Roughness

– Source: swissALTI^{3D} for elevation
– Sampling date: 2012/08/19 [Swisstopo n.d.(a)]
– Spatial resolution: 2m
– Maximal error per data source [Swisstopo n.d.(a)]:
    – up to 2000m a.s.l., $\pm$0.5m in all dimensions
    – above 2000m a.s.l., $\pm$3m
– Maximal error[4]: $\pm$3m
– Computation method: Various methods to calculate the roughness exist [Cooley 2016]. I decided to use the third method of Cooley 2016, which required a DEM raster on which the following focal statistics were applied: mean, minimum and maximum. I used the function *Focal Statistics* in ArcMap with a circle of radius 25m as neighbourhood.

$$roughness = \frac{mean - DEM}{maximal - minimal} \tag{3}$$

Then, I computed the roughness using the *Raster Calculator* in ArcMap and the formula 3.
– Remarks: First, I excluded the variable roughness percentile from analysis because of its strong correlation with aspect percentile, see Figure 66. Second, creating the roughness raster reminded me that raster calculations should be conducted - if possible - on inputs slightly bigger than the area of interest. The computation of the roughness required an elevation raster as input. To save computation time, I used the elevation raster which was cropped to the study area. Unfortunately, I realised too late that this introduced edge effects, see the red borders in Figure 55. Luckily, no data points were affected by these edge effects. Consequently, I decided to ignore this inaccuracy. However, the resulting edge effects could have been removed by cropping if the raster was bigger than the area of interest.
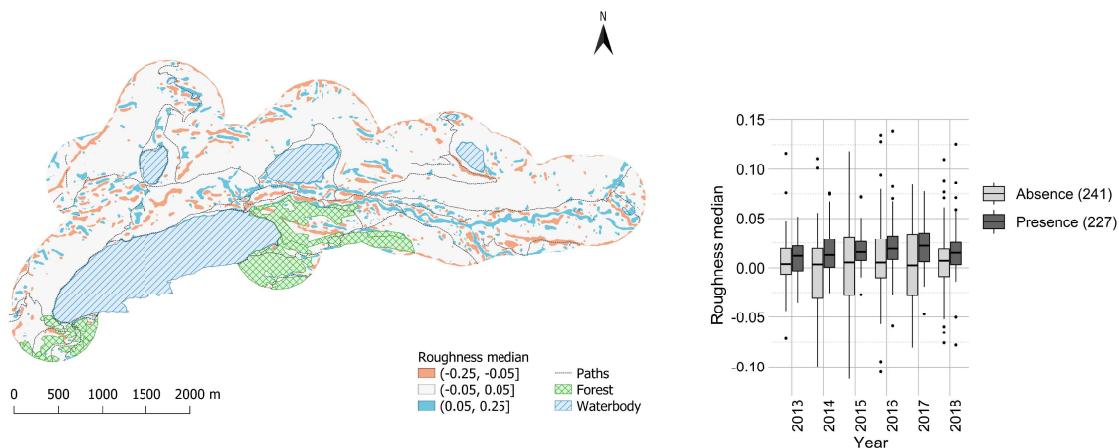
(a) The applied median focal statistic smoothed the original strongly, see Figure 55. Nevertheless, the valley of the Murinascia Grande and the most pronounced ridges emerge.

(b) The absence points vary largely compared to the presence points. Yet, the variation and the roughness itself are miniscule and confronts my expectation and the literature.

Figure 11: The median of the roughness could theoretically range between -1 and +1. The range in the Val Piora covers a mere fourth.

2.3.7 Normalized difference vegetation index (NDVI)

– Source: SWISSIMAGE false coloured infrared orthophotomosaic (FCIR) with the channels PAN, NIR, R, G and B
– Sampling date: 2015/08/30 [Swisstopo n.d.(b)]
– Spatial resolution: 0.5m scaled to 0.25m [Swisstopo n.d.(b)]
– Maximal error per data source: ground resolution ±5m in uneven areas [Swisstopo n.d.(b)]
– Maximal error[4]: ±5m
– Computation method: The spectral resolution for the channels near infrared (NIR), red and green is 16bit. I calculated the NDVI using the function *Focal Statistics* in ArcMap with a circle of radius 25m as neighbourhood.

$$NDVI = \frac{NIR - Red}{NIR + Red} \tag{4}$$

The calculation of the NDVI requires the channels NIR and red, see formula 4 [Lillesand et al. 2008].
– Remarks: The nest 2015_CH_Piora_A09 was exactly on the border between two raster tiles. Thus, I added the value found in its immediate surrounding manually.

.
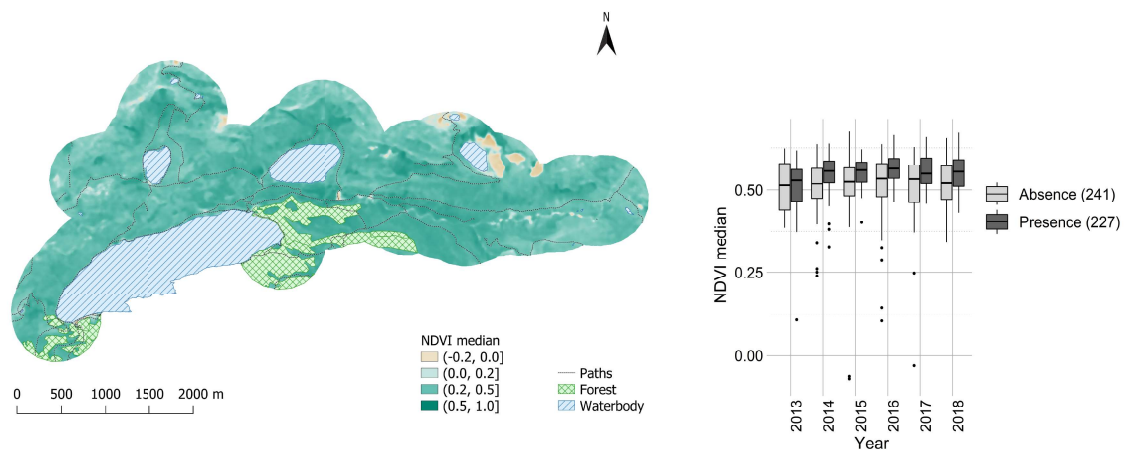
(a) The differentiation between vegetation (green) and bare soil or rocks (brown) works well, e.g. cairns were visible in the original, see Figure 57. The N-exposed slope eastern of the Murinascia Grande is densely vegetated and appears slightly darker. However, most of the study area features values between 0.5 and 0.6 suggesting rather dense and healthy vegetation, which might correspond to the late sampling date.

(b) The presence points range less and are at the top end of the variation of the absence points, which challenges my expectation. Presence and absence points exist mostly in vegetated areas and mirror the value distribution of the map.

Figure 12: The median of the NDVI could theoretically range between -1 and +1. The range in the study area is right shifted.
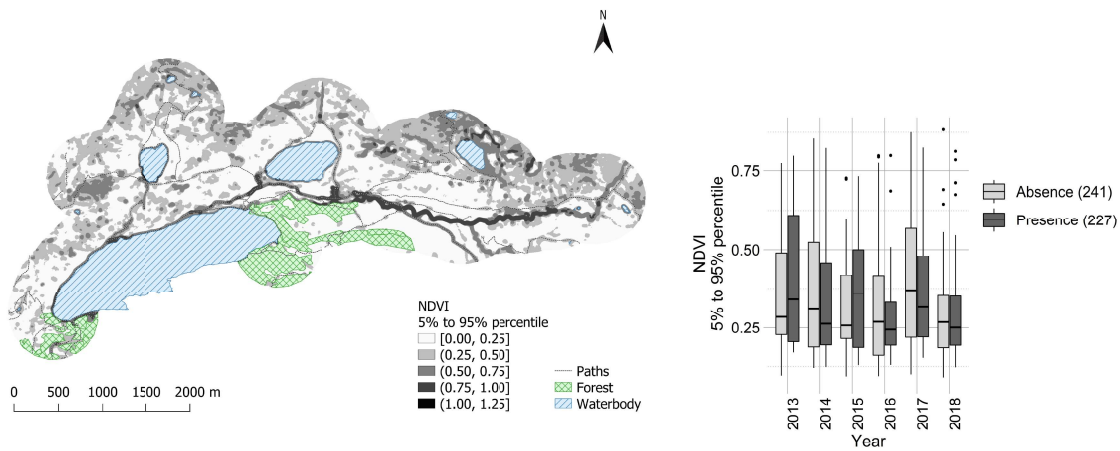
(a) The areas around paths, the Murinascia Grande and the cliffs in the north display variation. These areas contain rocks or bare ground.

(b) Most of the presence and absence points vary rather slightly given the possible range of values.

Figure 13: The 90% percentile distance of the NDVI could theoretically range between 0 and 2. However, the NDVI in the flatter part of the study area is rather uniform and might prevent the expected heterogeneity in land cover. This might indicate that northern wheatears favour flat slopes over sparsely vegetated areas (as long as suitable hunting grounds are near).

2.3.8 Cumulated distance to the three nearest nests

– Source: swissALTI^{3D}
– Sampling date:
  – 2012/08/19 for swissALTI^{3D} [Swisstopo n.d.(a)]
  – variable for location
– Spatial resolution: 2m for swissALTI^{3D}
– Maximal error per data source:
  – swissALTI^{3D} [Swisstopo n.d.(a)]:
    – up to 2000m a.s.l., $\pm$0.5m in all dimensions
    – above 2000m a.s.l., $\pm$3m
  – location: $\pm$10.5m, see 6.2.1
– Maximal error: $\pm$66m
– Computation method:

$$d = \sqrt{\Delta x^2 + \Delta y^2 + \Delta z^2} \tag{5}$$

First, I calculated a distance table for each point (presence or absence) to every presence point of the respective year with formula 5, where x is the easting, y is the northing and z is the elevation. Then, I applied the following steps to every point: (1) select the distances starting at point p1 and ending at any nest (except p1), (2) sort the selected distances ascending, (3) chose the first three distances and (4) build their sum.

– Remarks: There was a difference between the 2D and 3D cumulated distances (straight line). I chose 3D as I assume that this method approximates the behaviour of a bird.
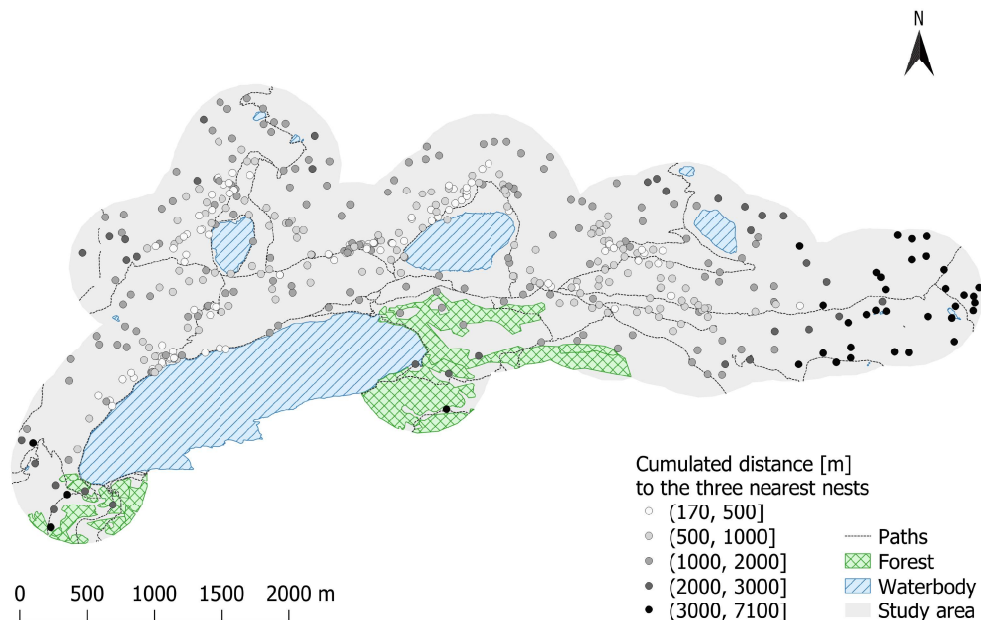

Figure 14: Small cumulated distances indicate clusters. There are four especially clustered areas: north-west shore of Lago Ritom, north-west shore of Lago Cadagno, north-west shore of Lago di Tom and the area between the Capanna Cadagno and Lago di Dentro. The cumulated distance to three nearest nests in the study area ranges between 173m and 7100m.
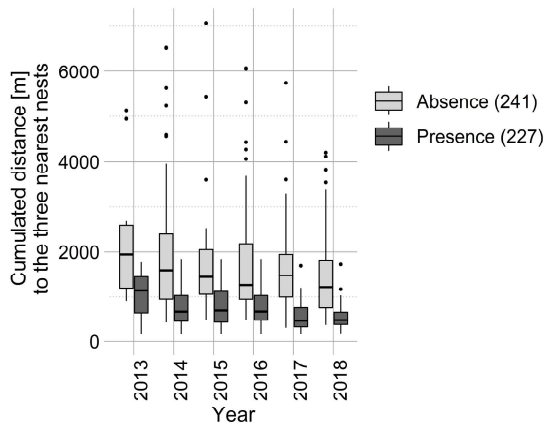
Figure 15: The presence points are generally closer to their three nearest neighbours.

Year	Minimum	Mean	Maximum
2013	15.32	338.79	839.37
2014	35.09	150.51	522.25
2015	35.63	169.30	583.10
2016	22.31	121.80	443.09
2017	26.01	152.59	599.89
2018	21.51	135.98	306.74

Table 3: The distance to the nearest nest per year ranges between 15m and 839m.

I computed the minimal distance between two nests to compare the found minimal distance to the existing literature and review the chosen radius for the definition of the nest site. An intermediate result was the distance to each other nest. From this, I extracted the distance to the nearest nest in the Val Piora per year, see Table 3.

The literature value of the distance to the nearest nest in the Swiss Alps varied between "often barely 100m" and "nests are not uncommonly closer than 200m" [Corti and Melcher 1958; Glutz von Blotzheim 1962].

2.3.9 Distance to the nearest path

– Source:
  – swissALTI^{3D} for elevation
  – swissTLM3D Strassen for path
– Sampling date:
  – 2012/08/19 for swissALTI^{3D} [Swisstopo n.d.(a)]
  – 2015 for swissTLM3D [Swisstopo n.d.(c)]
– Spatial resolution: 2m for swissALTI^{3D}
– Maximal error per data source:
  – swissALTI^{3D} [Swisstopo n.d.(a)]:
    – up to 2000m a.s.l., $\pm$0.5m in all dimensions
    – above 2000m a.s.l., $\pm$3m
  – location: $\pm$10.5m, see 6.2.1
  – swissTLM3D: $\pm$1.5m [Swisstopo n.d.(c)]
– Maximal error[4]: $\pm$12.5m
– Computation method: I computed the spatial distance to the nearest path in ArcMap with the tool *path distance* and the above DEM as surface raster.
– Remarks: The study area was computed with a buffer of 500m around certain paths. However, the map shows values up to 750m. This discrepancy originates in the buffer being calculated as 2D straight line and the distance to the path as ground distance on above DEM.
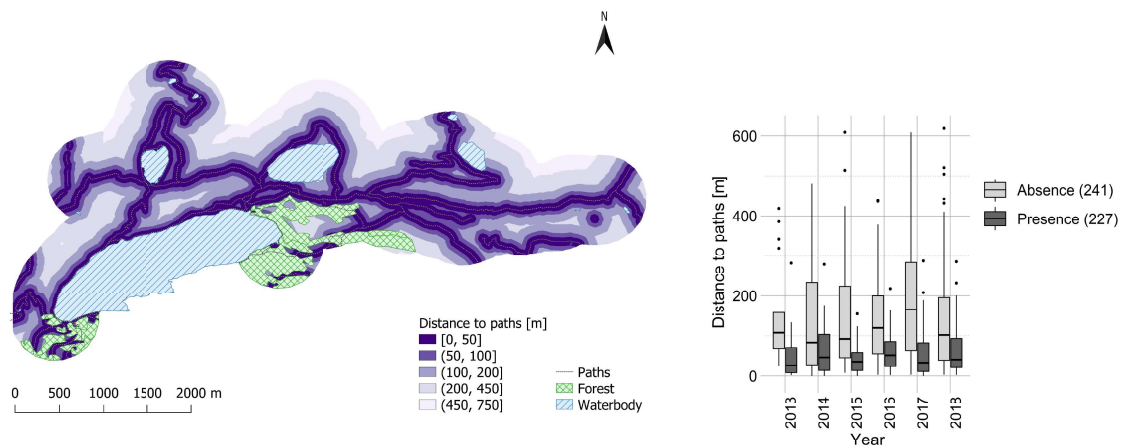
(a) Paths cross especially the flat part of the study area.

(b) The presence points scatter generally closer around the paths. The northern wheatears nest hardly ever further than 100m from the nearest path.

Figure 16: The distance to the nearest path in the study area ranges between 0m and 750m.

2.3.10 Nest densities

I computed the nest density per year to compare the local densities found in the Val Piora not only to the literature values but also to the value of a uniform distribution. I wondered how many nests would fit in 10ha if they were distributed evenly. Thus, I constructed a square with area 10ha (side length = 316m). I used the average distance to the three nearest nests (721m) to construct a uniform distribution in GeoGebra, see Figure 17. Under the assumption of uniformity, two nests were on average 240m apart.
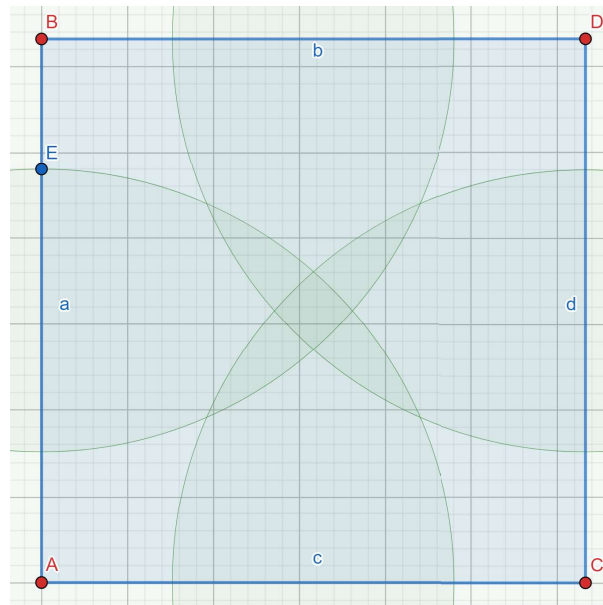

Figure 17: A square with side length 316m approximates an area of 10ha. The distance between two nests equals 240m ($\overline{AE}$), which is the radius of the displayed circles. A, B, C and D are nests, a fifth nest could not be fit into the square.

I showed empirically that ~4 nests would fit in a square of 10ha. Incidentally, a circle of area 10ha could also contain four uniformly distributed nests. However, the nests in the study area disclosed no uniform distribution, see Figure 4.

I wondered what the number of nests in 10h around a nest was in the Val Piora per year. Thus, I created a buffer around each nest with radius 178m resulting in circles of area 10ha. Then, I counted the number of nests within each individual buffer per year, see Figure 72. I extracted the local minimum, local mean and local maximum density [nests/10ha] from the layer statistics, see Table 7.

The density of nests within the entire study area was assumedly smaller as the marginal area was free of nests, see Figure 4. Nonetheless, I compute this measure by building the ratio of the number of nests per year and the area of the study area in 10ha. Table 7 summarises the results in the column Overall [nests/10ha]. I refrained from further clustering analysis because of shortage of time.

2.3.11 Distance Measures

One of the first questions to answer was how to define the study area. I decided to include the area which was closer than 500m to the patrouilled paths [Meier 2018]. However, a distance of 500m can be computed in at least three ways. The three distance measures straight line, slant range and ground distance compute the distance differently and thus differ in their outcome, especially in a mountainous area such as the Val Piora [Bürki 2014]. For each of these three methods, I shortly summarise the idea of computation and where I used it.

The straight line computed the Euclidean distance in a planar environment, considering only easting and northing. It was the standard distance for buffers and easily interchangeable between the platforms QGIS and ArcGIS. This measure was used to buffer the the paths in the creation of the study area and to buffer the presence points to calculate the pseudo-absence points.

The slant range calculated the Euclidean distance in a three-dimensional environment, considering easting, northing and elevation, [Lillesand et al. 2008]. Distances among a small number of points were effortlessly analysed. This method approximated the trajectories of flying animals as long as no ridges blocked the direct path. Therefore, I used this for the computation of the distance to the nearest nest and for the cumulated distance to the three nearest nests.

The ground distance measured the path length on top of a DEM as if somebody walked from the start to the destination. This measure included the additional distance of obstacles, such as a hill or valley. I used this to compute the distance to the nearest path. Originally, the distance to the nearest path was intended as approximation for disturbances of hikers [Bötsch, Tablado, and Jenni 2017; Bötsch, Tablado, Scherl, et al. 2018; Finney et al. 2005]. Hikers might occasionally leave the marked paths but they are bound to the ground.

2.4 The snow coverage algorithm

2.4.1 Quick overview

Northern wheatears have been observed to prefer nest sites which are snow-free by the end of May [Wartmann 1985]. For further investigation, the Swiss Ornithological Institute (SOI) installed three cameras which have been recording the summer seasons in Val Piora between 2013 and 2018. Here, I present the developed algorithm which extracts the course of the snow coverage for a nest site during a summer season from this footage. The key inputs were a *Photo*, see Figure 18a, taken by one of the installed cameras and the corresponding scene in a 3D digital elevation model (DEM), where the nest site is marked in red, see Figure 18b.

(a) *2-Photo*: taken by the camera monitoring the Lago Ritom in 2015.   (b) *1-Model*: example of a nest site (red) at the Lago Ritom on the DEM.

Figure 18: The snow coverage is only analysed for the region marked by red pixels in the *1-Model*.

The basic workflow of the developed algorithm was: (1) Prepare the necessary data, see 2.4.3. (2) Adjust the *2-Photo* (Figure 18a) such that it is congruent with the *1-Model* (Figure 18b), see 2.4.4 and 2.4.5. (3) Determine the position of the nest site (red pixels) in the *1-Model* and address the corresponding pixels in the *2-Photo*; count these pixels; calculate the ratio between pixels covered in snow and the total number of pixels of the red area. (4) Apply the third step on every suitable *2-Photo* and thus analyse the snow coverage over the course of one season, see 2.4.6.
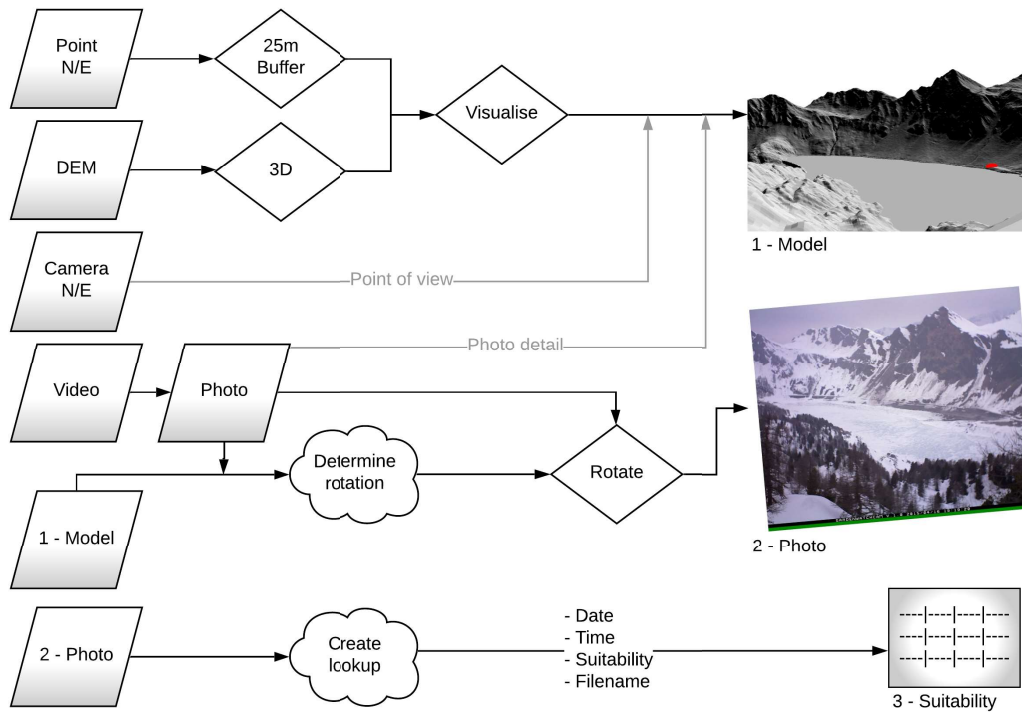
Figure 19: The exact steps for the data preparation are explained in 2.4.3. Camera (N/E) equals the recorded northing/easting of the position in the Swiss Grid (LV03). **Please note:** entries with a number, e.g. *1-Model*, are used in the text below to link to these flow charts.
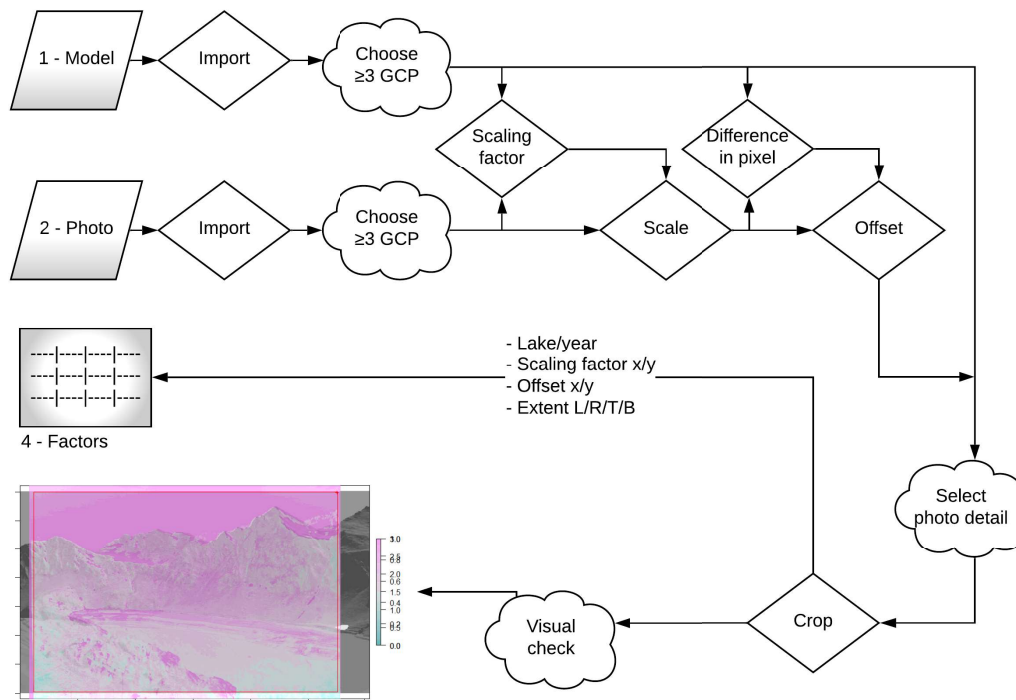
Figure 20: The exact steps of overlay are explained in 2.4.4. Extent L/R/T/B equals the left/right/top/bottom margin of the cropped *2-Photo*. The ground control points (GCPs) mark distinct features on *1-Model* and *2-Photo*. The scaling factor x/y and offset x/y describe the horizontal/vertical case individually.
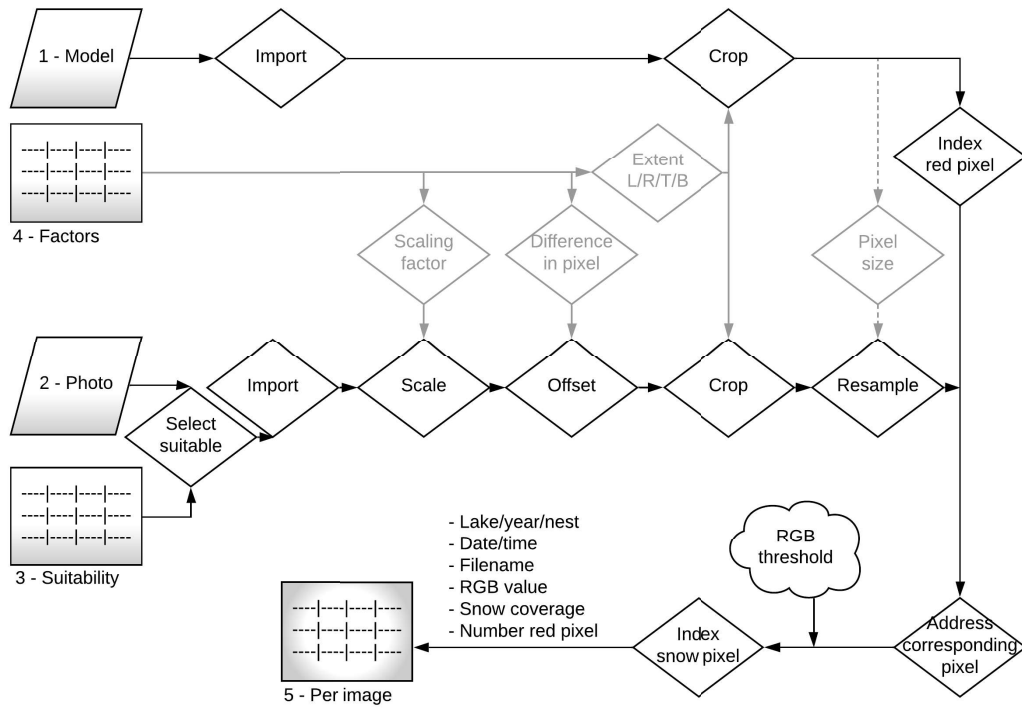
Figure 21: The exact steps of snow coverage per nest site and image are explained in 2.4.5.
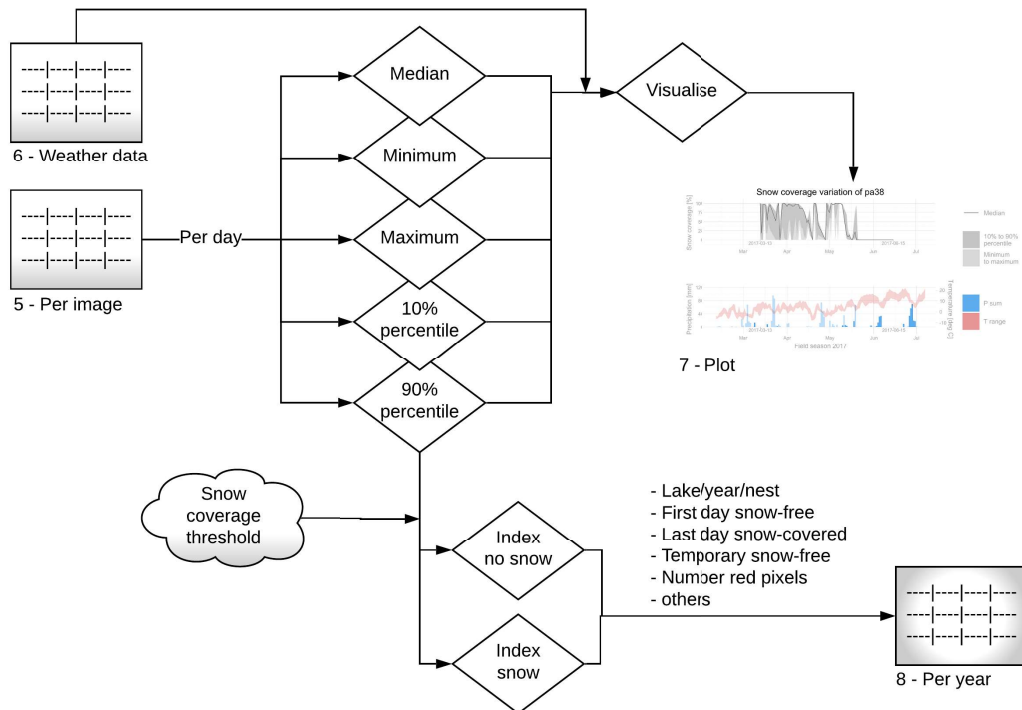
Figure 22: The exact steps of snow coverage per nest site and season are explained in 2.4.6. The median, minimum, maximum, 10% percentile and 90% percentile were calculated from all computed snow coverage values for any day.

2.4.2 Introduction and setup

Three cameras recorded the surroundings of Lago Ritom, Lago Cadagno and Lago di Tom each summer season between 2013 and 2018. They were installed in early spring and took a picture every hour, until the battery was drained or until they were collected at the end of summer. The cameras were mounted to document the areas with anticipated nests. However, the topography and vegetation prohibited to record all nest sites; Some nest sites were only partially visible. An area-wide documentation is not possible with the current setup. Nonetheless, this approach produced thousands of *Photos*, which demanded an automated analysis method. Unfortunately, I found no available georeferencing tool, especially not in R, which could automatically georeference 2D-photos on a 3D-DEM. Existing image correlation software such as CIAS [5] lack the possibility to automatically treat multiple pictures. Hence, I wrote this algorithm which not only localises the nest site on the *2-Photo* but also analyses its snow coverage.

2.4.3 Preparation

I started with cutting the time-lapse video footage into *Photos*. Simultaneously, each photo was manually categorised depending on the meteorological condition. By this procedure all pictures with signs of clouds, precipitation, haze, partial shadow or strong overexposure could later be excluded from the analysis. *3-Suitability* contained per *2-Photo*: filename, date, time and suitability for analysis. The preparation produced three outputs *1-Model, 2-Photo* and *3-Suitability*. Each was prepared in a separate process, see Figure 19.

The creation of *1-Model* required the coordinates of all presence and absence points (*Point N/E*), the *DEM*, the coordinates of the camera (*Camera N/E*) and a reference *Photo*. First, I created a 25m buffer around each point (presence or absence) in ArcScene to represent the nest sites. A nest site is the area within 25m around a presence point or an absence point. Then, I 3D-visualised the DEM with the buffers drawn in red on top. Next, I navigated in this 3D-visualisation to the camera location to recreate the point of view of the camera and to focus on the detail depicted in the reference *Photo*. Finally, I exported the just created scene depicting each nest site individually, see Figure 18b.

To obtain *2-Photo*, I rotated each *Photo*, as their orientation was slightly rotated compared to the *1-Model*. To determine the rotation of the *2-Photo*, I laid a semi-transparent reference *Photo* over the *1-Model* in Microsoft Word. Then, I rotated the reference *Photo* until its horizon line matched the horizon line of the *1-Model*, see Figure 18a. Afterwards, a short R-script rotated all *Photos* taken by the same camera in that year.

---

[5] https://www.mn.uio.no/geo/english/research/projects/icemass/cias/#features

2.4.4 Overlay

This step adapted the images such that the *2-Photo* congruently covered the *1-Model* afterwards. A congruent covering is achieved when a mountaintop X is located at the same position in *2-Photo* and in *1-Model*, see Figure 20. I applied this method once per lake and year.

To begin, I imported the *2-Photo* and the *1-Model*. Then, I chose at least three ground control points (GCPs) on the *1-Model* and the same GCPs on the *2-Photo* by mouse clicks on those locations, see Figure 23. GCPs typically are "physical points on the ground whose ground positions are known with respect to some horizontal coordinate system and/or vertical datum" [Lillesand et al. 2008]. GCPs could be chosen on a distinct landmark such as a mountaintop, building, road junction etc. The coordinates of the here used GCPs are unknown. Nonetheless, their relative position on the *2-Photo* and the *1-Model* sufficed to bring them, and consequently the entire detail, to congruency. Absolute georeferencing was not needed.
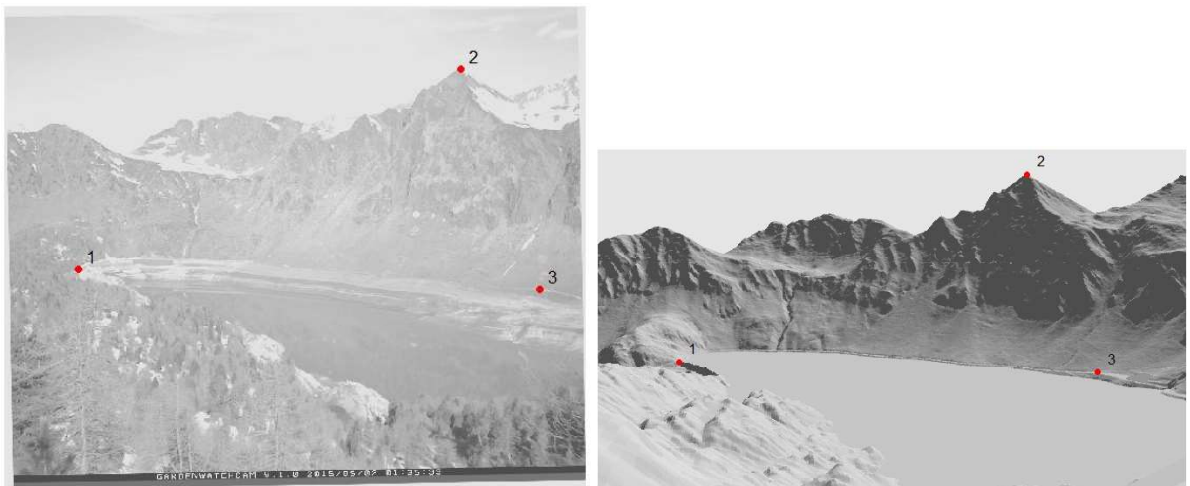
(a) This *2-Photo* was taken by the camera monitoring the Lago Ritom in 2015. The three dots are the chosen GCPs that match the GCPs of the *1-Model*, see Figure 23b. GCP1 is the end of the dam, GCP2 is a mountain top and GCP3 is a distinct point along the path.

(b) This *1-Model* depicts the 3D-visualised DEM and the same three GCPs of Figure 23a. The Lago Ritom is dammed up to the maximum. The GCPs are internally indexed from left to right.

Figure 23: The detail and the point of view of the *2-Photo* and the *1-Model* are already nearly congruent.

The GCPs need to spread over the entire picture, with some GCPs chosen in the vicinity of nest sites. I explain the following steps using an abstract representation of three GCPs:
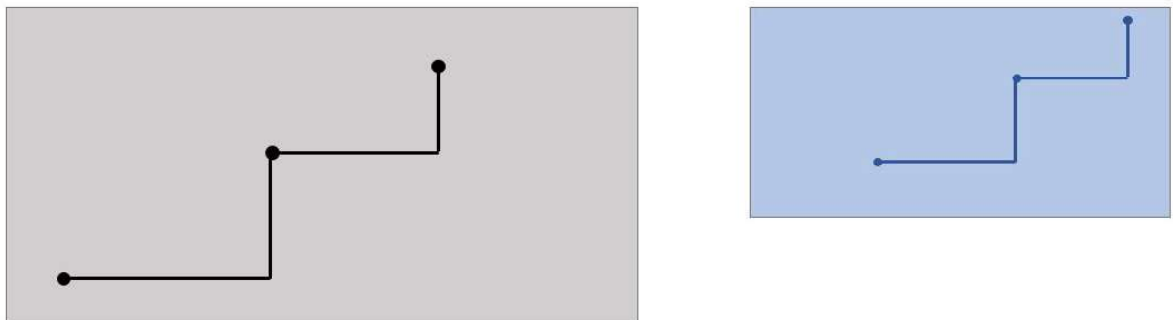

Figure 24: The *2-Photo* (left) and the *1-Model* (right) feature each three GCPs. The connecting lines are only a visual aid.

After choosing the GCPs, I calculated the scaling factors. A scaling factor described to what extent a

2-Photo had to be stretched or compressed so that the same distance laid between corresponding GCPs in the *1-Model* and in the *2-Photo*. There was a scaling factor in x-direction and one in y-direction. I started by calculating the distance between each GCP and the next GCP, see formula (6).

$$model\_diff\_x = dplyr :: lead(model\_gcp\_x) - model\_gcp\_x \tag{6}$$

$$photo\_diff\_x = dplyr :: lead(photo\_gcp\_x) - photo\_gcp\_x \tag{7}$$

Let model_gcp_x be an array of the x-coordinates of the GCPs of the *1-Model* and photo_gcp_x the same for the *2-Photos*, see formulas (6) and (7). Next, I divided the calculated differences of the *1-Model* by the differences of the *2-Photo*. This created an array of scaling factors. The mean of this array resulted in an overall scaling factor, see formula (8).

$$scale\_x = mean(model\_diff\_x / photo\_diff\_x, na.rm = T) \tag{8}$$

I repeated this calculations for the scaling factor in y-direction. Subsequently, I scaled the *2-Photo* accordingly, including the position of the GCPs, see Figure 25.

Figure 25: The now scaled *2-Photo* (left) displays the same distance between two GCPs as the *1-Model* (right).

If I laid the so scaled *2-Photo* over the *1-Model*, it would look like Figure 26.
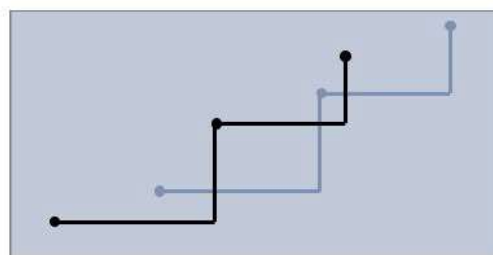
Figure 26: The scaled *2-Photo* lies over the *1-Model*. Yet, the GCPs are not congruent as the *2-Photo* detail is slightly shifted compared to the *1-Model*.

To rectify the above shift between the pictures, I calculated the number of pixels between the scaled GCPs of the *2-Photo* and the corresponding GCPs of the *1-Model*. This resulted in an array of local offsets. The mean of this array equals an overall offset in x-direction, see formula 9.

$$offset\_x = mean(model\_gcp\_x - photo\_gcp\_x\_resized) \tag{9}$$

Let photo_gcp_x_resized be the scaled GCPs of the *2-Photo*. I repeated this in the y-direction. Subsequently, I shifted the entire *2-Photo* by these offsets. If we laid the *2-Photo* over the *1-Model* now, it would look like Figure 27.
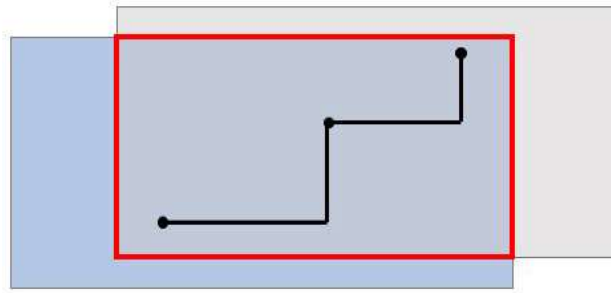
Figure 27: The scaled and offset *2-Photo* (grey) covers congruently the *1-Model* (blue). The red rectangle marks the shared detail.

The application of these steps to a real *2-Photo* and *1-Model* resulted in Figure 28. The best fit is generally near the GCPs.
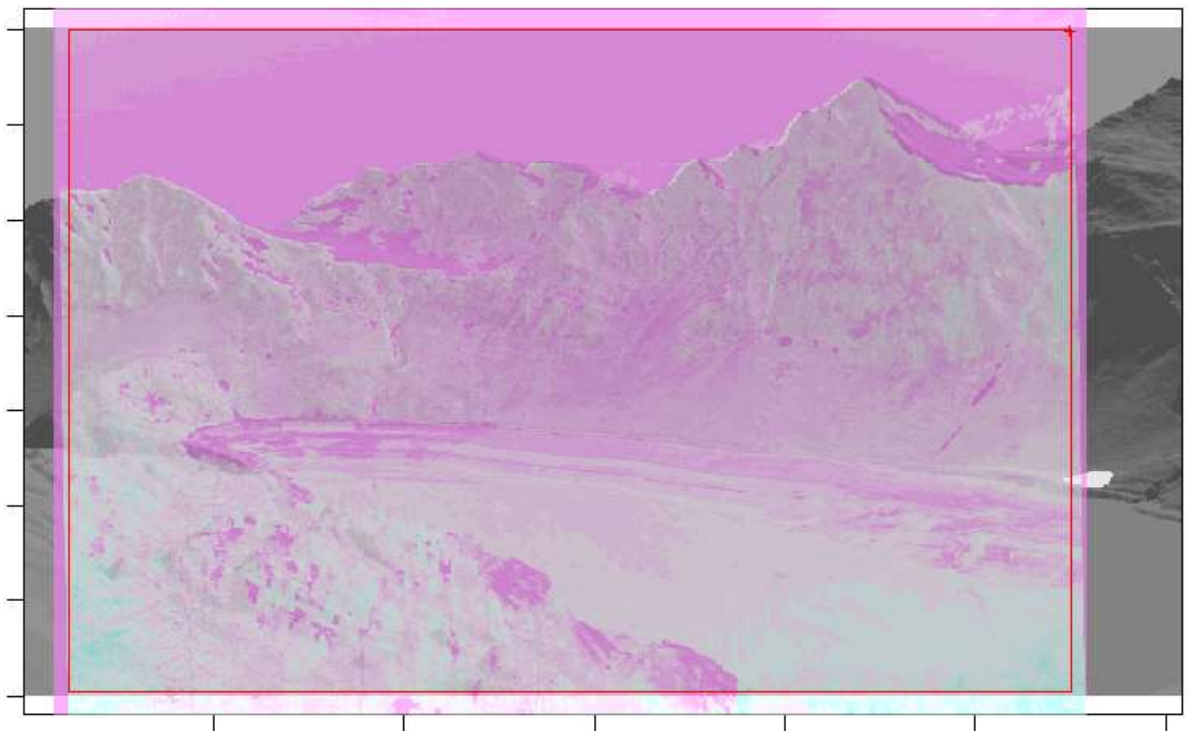
Figure 28: The scaled and offset *2-Photo* (blue-pink) lies semi-transparent over the *1-Model* (black-and-white). It allowed a visual check of the quality of the overlay. Here, it is not a perfect fit as the light border at the horizon indicates. However, the road along the Lago Ritom is not visible twice which would provide clear evidence of an imprecise fit. The nest site (white) lies barely in the shared detail, which is marked with the red rectangle.

I cropped the *2-Photo* and *1-Model* to the shared detail. Thus, future analysis will only affect pixels which are available in both pictures. Mouse clicks on the lower left and top right corner of the shared detail cropped the *2-Photo* and the *1-Model*. Then, I stored the camera location, year, scaling factor in x- and y-direction, offset in x- and y-direction and the left (L), right (R), top (T) and bottom (B) extent of the shared detail in *4-Factors*. Finally, I turned this workflow into a function and applied it to every possible lake and year combination.

2.4.5 Snow coverage per nest site and image

This function analysed the snow coverage on one *2-Photo* for one nest site, see Figure 21. The taken *Photos* were recorded in the bands red, green and blue (RGB). I considered a pixel as snow if it was whiter than a certain RGB threshold. To determine the RGB threshold, I manually sampled patches of snow on *Photos* taken at the Lago di Tom, see 6.2.13; Snow was usually whiter than RGB(0.6, 0.6, 0.75). Another approach would have been the use of an RGB normalised difference snow index (NDSI) [Hinkler et al. 2002], see 6.3.3. The NDSI discovers snow and is based on the difference in spectral behaviour of snow, vegetation and soil [Hinkler et al. 2002]. The high reflectance of snow in near-infrared (NIR) and in mid-infrared (MIR) rapidly decreases while bare soil and vegetation behave differently [Duguay and LeDrew 2006]. However, RGB photos lack a NIR or MIR band which prohibits the computation of the NDSI. To approximate an NDSI using only RGB photos, Hinkler et al. 2002 introduced two methods to compute an RGB NDSI. Their general idea of both methods is to replace the missing MIR band by a combination of RGB bands. In their first approach, the MIR-replacement has taken the overall level of brightness into account and thus corrected for partially darkened photos. Such partially darkened photos originated in partial cloud covering. In their second approach, they only considered the mean RGB value for the MIR-replacement which sufficed in most conditions. Their empirical threshold value for snow has equalled 0.6. [Hinkler et al. 2002]

The actual analysis began with *2-Photo* and *3-Suitability*. I listed all *2-Photos* taken by a specific camera in a given year, but I imported only those *2-Photos* which were labelled suitable in *3-Suitability*. The following process chain was applied to each combination of these *2-Photos* and *1-Models*. I describe the proceeding on one suitable *2-Photo* and one *1-Model*. The lookup table *4-Factors* contained lake, year, scaling factor, offset and extent. Its composite key was lake or camera location and year. At the begin, I extracted the scaling factors of that camera and year from *4-Factors* and scaled the *2-Photo* accordingly. Then, I imported the differences in pixels from *4-Factors* and offset the *2-Photo* correspondingly, see Figures 26 to 27. Next, I imported the *1-Model*, loaded the extent of the shared detail from *4-Factors* and cropped the *1-Model* and the *2-Photo* accordingly. Afterwards, the *1-Model* and the *2-Photo* covered the same extent and were visually congruent.

However, for indexing I required not only visual but pixel-wise congruence. Thus, I resampled the *2-Photo*, i.e. changed its raster resolution, using the *raster::resample* function and the *1-Model* as reference resolution. Hence, the *2-Photo* and the *1-Model* shared extent, visual and pixel-wise congruency. The mountaintop X is located now on the same pixel in the *2-Photo* as in the *1-Model*, not only on the same position. Then, I searched for all red pixel in the otherwise black-and-white *1-Model* and created an index of the location of their pixels. Since the *2-Photo* and the *1-Model* were pixel-wise congruent, I addressed the pixel of the nest site in the *2-Photo* using the just created red-pixel-index.

Finally, the snow coverage analysis of this addressed nest site commenced by counting the number of pixels of the red-pixel-index which described the size of the nest site. Then, I selected those pixels of the red-pixel-index which were whiter than the previously empirically determined RGB threshold and counted the number of pixels in this selection of snow pixels. The snow coverage equalled the number of pixels considered as snow in relation to the number of pixels of the investigated nest site and thus indicated the ratio of the visible surface of the nest site covered in snow, see formula (10).

$$snow\ coverage = snow\ pixels\ /\ red\ pixels \tag{10}$$

I stored this snow coverage in a spreadsheet together with the metadata: lake (camera location), year, nest, date, time, filename, used RGB threshold and the number of red pixels. Eventually, I turned these steps into a function and applied it to every combination of *2-Photo* and *1-Model*. The final data set contained one entry per possible combination of *2-Photo* and *1-Model* for the chosen lake and year.

The above algorithm performed accurately under neutral illumination. However, extreme illumination conditions caused inaccurate performance. Thus, I removed the majority of overexposed *2-Photos* to improve the accuracy of the result. After this additional selection, a snow-free day surrounded by snow-covered days could indicate a or multiple *2-Photos* which were too dark or bluish for correct snow detection. This error is visible at the end of March in Figure 29 and affected the first snow-free day. On the other hand, an isolated peak in the snow coverage without preceding light blue precipitation column in the climatogram hinted at a or multiple overexposed *2-Photos* which prevented correct surface distinction. This mistake is not present in Figure 29 and affected mostly the last snow-covered day. Consequently, the temporal accuracy is undefined but large.

The statistical measure to determine the daily snow coverage was computed fast and can be readily changed to a different measure. The final algorithm used the 90% percentile as such measure. The chosen measure influenced the computed snow coverage largely, see the difference of the minimum, 10% percentile, median, 90% percentile and maximum in Figure 29. During the development of the algorithm, I scrutinised the measures 90% percentile per day, maximum per day and median over three days. The maximum per day was computed accordingly to the 90% percentile per day, just with the maximum function. The median over three days included also the preceding and the following day into the analysis. The such computed median could not be calculated for days where a neighbouring day was missing *2-Photos*. The median over three days was discarded because it strongly smooths the snow coverage, the smoothing affects especially the snow-melting period and it loses every fourth day of data because it can not be computed around days without suitable *2-Photos*. The maximum per day was discarded although there was no loss of data and fresh snow events were recognised. However, it was extremely affected by overexposure which was a constant issue with the used camera. Thus, I chose the 90% percentile as there was no loss of data and its slight smoothing could filter the most severe effects of overexposure, compare the daily maximum to the daily 90% percentile in Figures 74 to 76.

The accuracy assessment of the detected snow coverage per nest site and image was accomplished with a visual comparison to the *2-Photos* and a juxtaposition of double-monitored nest sites. The qualitative visual comparison of the calculated snow coverage with the analysed *2-Photo* revealed that most nest samples appeared to be within an estimated 10% of the specified coverage.

There was a single nest which was photographed by two cameras simultaneously. I assessed the accuracy of the snow coverage algorithm by comparing the results derived from the different cameras for this nest site.

2.4.6 Snow coverage per nest site and season

I intended this method to describe the course of snow coverage for one nest site in one season with statistical measures and a visual comparison to the weather condition at that time, see Figure 22.
I started by extracting all computed snow coverage values and their timestamp of a specific nest from *5-Per image*, see Figure 19. Usually, multiple entries existed per day. Then, I computed the minimum, the 10% percentile, the median, the 90% percentile and the maximum of the snow coverage for each day and visualised these measures together with the precipitation and temperature data, see Figure 29.
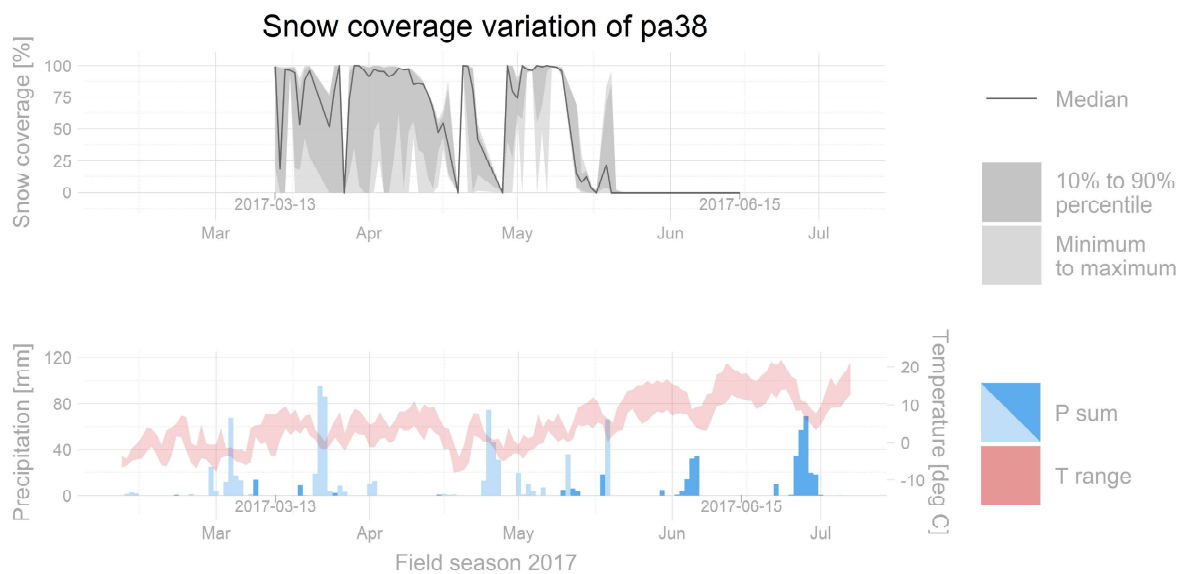

Figure 29: This plot contains two subplots which both depict the course of the field season (here: 2017). The two dates on the x-axis mark the start and end date of the recording. The top subplot discloses the daily snow coverage. For the same period, the bottom subplot displays the precipitation sum per day (P sum, left axis) and range between the minimal and maximal daily temperature (T range, right axis). The precipitation bars are dark blue if rain was expected and light blue if snowfall was assumed, i.e. when the minimal temperature was below 1° C. This plot assisted the accuracy assessment.

I only considered the 90% percentile of the daily snow coverage for further analysis, e.g. comparison with the snow threshold of 10%. I defined an area as snow-free if less than 10% of the pixels were identified as snow; otherwise the area was defined as snow-covered. Then, I calculated (1) the first snow-free day, which equalled the date in days passed since 1st January of the first recorded snow-free day, (2) the last snow-covered day, which was the date in days passed since 1st January of the last recorded snow-covered day, (3) the number of temporary snow-free days counted the number of snow-free days between the first snow-free day and the last snow-covered day, (4) the recorded total number of days with snow respectively without snow, (5) if the nest site was snow-free from start (true or false) and (6) if there was snow after the first snow-free day (true or false). Subsequently, I stored these values and the metadata lake, year, nest, number of red pixels, calculation method and the RGB threshold value in a spreadsheet. Before I decided to use the 90% percentile, I experimented also with the daily maximal snow coverage and the median snow coverage over 3 days, see 6.3.3. Finally, I applied this method to all nest sites and added the obtained data to the data created in GIS, see 2.4.7.
The accuracy of the two outputs (*7-Plot* and *8-Per year*, see Figure 22) of this method was assessed with a visual analysis of the plot and the linear mixed effect models (LMEM). I qualitatively analysed various plots like Figure 29.

2.4.7 Further used snow coverage

– Source:
  – swissALTI^{3D} for elevation
  – GardenWatchCam for *Photos*
– Sampling date:
  – 2012/08/19 [Swisstopo n.d.(a)] for swissALTI^{3D}
  – various for *Photos*
– Spatial resolution: 2m for swissALTI^{3D}; undefined for *Photos*
– Maximal error per data source:
  – swissALTI^{3D} [Swisstopo n.d.(a)]:
    – up to 2000m a.s.l., $\pm$0.5m in all dimensions
    – above 2000m a.s.l., $\pm$3m
  – *Photos*: undefined
– Maximal error[5]: undefined, but large
– Computation method: see 2.4
– Remarks: This variable is spatially and value-wise the most inaccurate.
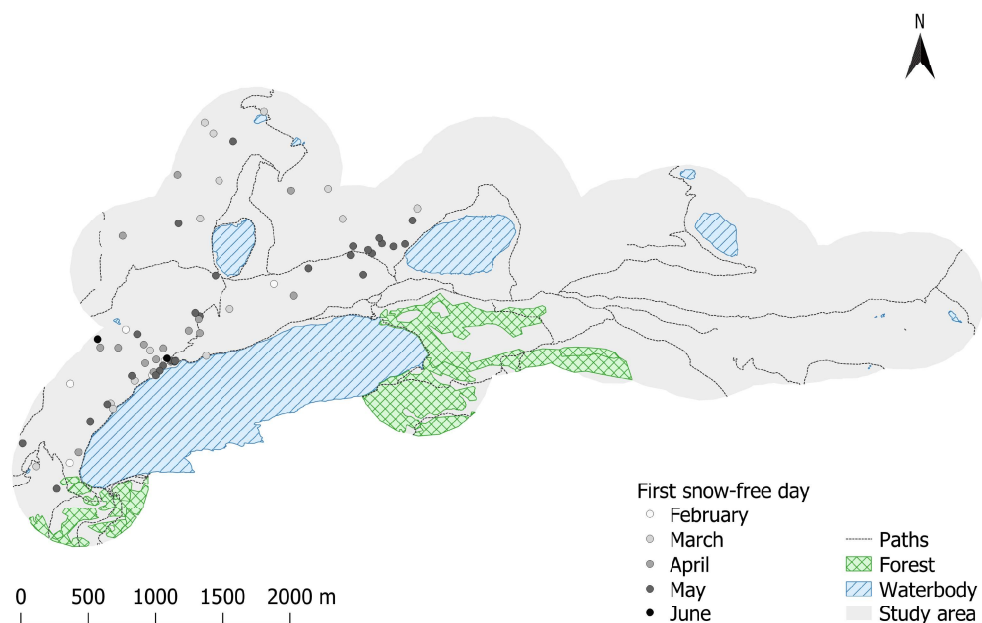
Figure 30: The first snow-free day falls between February and June. No clear spatial pattern is detectable.
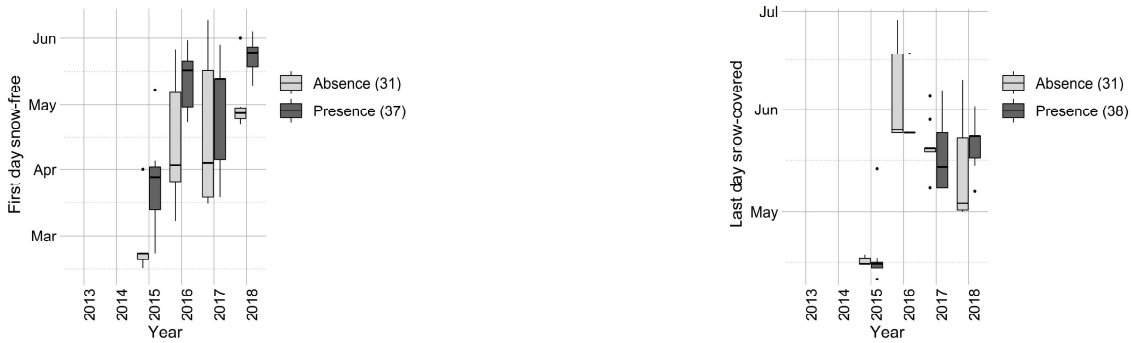
(a) Most locations are the first time snow-free until the end of April. Since the northern wheatears usually appear at the end of April or start of May, they arrive when most of Val Piora is already snow-free.

(b) The last day of snow falls often in May or June. However, this might not largely affect the breeding northern wheatears as the last snowfall often disappeared within one day and did not reach until the valley bottom.

Figure 31: The first snow-free day and the last snow-covered day show a large between-year variation. The explanatory power of these two variables may be limited because of the low sample size and the restricted field of view of the three cameras.
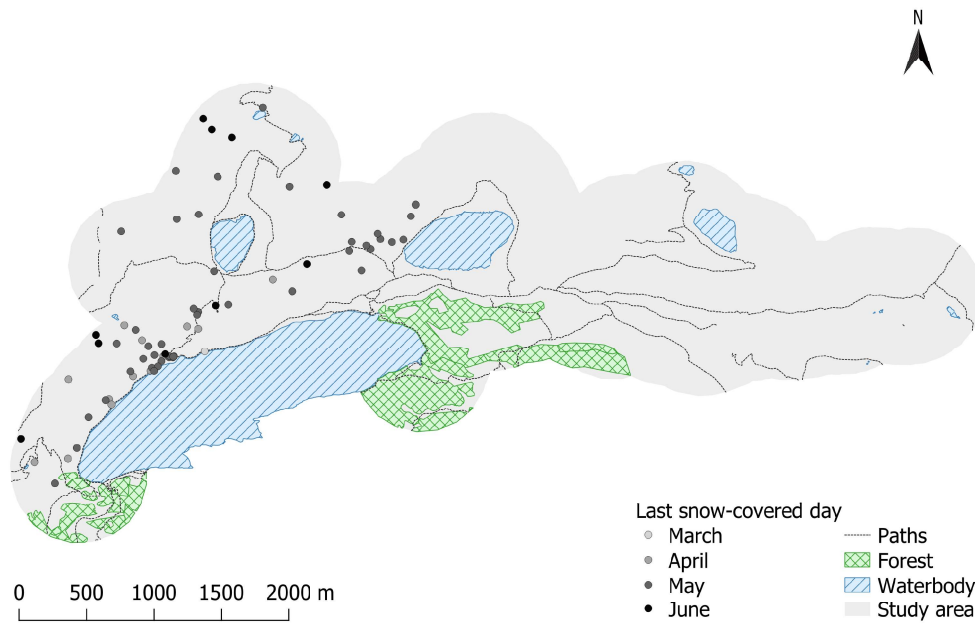
Figure 32: The last snow-covered day happens between March and June. Locations with the last snow in March or April exist only around the Lago Ritom.

2.5 The statistical models

I chose the Bayesian statistics because I wished to determine the likelihood of a presence point at a specific location given the analysed environmental factors [F. Korner-Nievergelt, Roth, et al. 2015]. I created three different models, which investigated various aspects of this question and three models scrutinizing the plausibility of the computed measures of the snow coverage over the course of the season.

2.5.1 Plausibility of the computed snow coverage variable

These models were created as proof of principle on the computed snow coverage variables (1) first snow-free day, (2) last snow-covered day and (3) number of temporary snow-free days. I included variables describing the environmental factors (1-7 and 12, see *Data* below) and methodological factors (8-11, see *Data* below). If the computed snow coverage variable was mostly predicted by environmental factors, this measure might describe the real snow coverage course; otherwise it was affected stronger by the chosen methods.

Data

I analysed 69 data points with computed snow coverage variables which laid outside of the Lago Ritom, see Table 4.

	2013	2014	2015	2016	2017	2018
Absence points	0	0	5	11	9	6
Presence points	0	0	8	16	8	6

Table 4: A summary of the number of the analysed presence and absence points per year.

I expected the following twelve variables to possibly affect the snow coverage variables: (1) Median of the NDVI; Different kinds of undergrounds might differ in their thermo-conductivities, which I expected to not only affect how quickly the snow stuck to the underground but also how long the snow remained. (2) 90% percentile distance of the NDVI; A change of the underground might influence the thermo-conductivities and might create edges (e.g. edges of the meadow) in which the snow was sheltered. (3) Median of the elevation; I expected lower altitudes to be covered in snow later and to thaw earlier because of a vertical thermal gradient. (4) Median of the aspect; South exposed slopes might to be snow-free earlier as the sun shines longest and strongest on unconcealed S-exposed features. (5) 90% percentile distance of the aspect; I assumed that the heterogeneous aspects originated from features such as hillocks, which protected the snow at some locations from direct sun radiation and thus from melting early. (6) Median of the slope; Plains might be early snow-covered and late snow-free because snow presumedly sticks easiest on flat terrains and potential avalanches might accumulate there creating long-lasting snow-cones. (7) Median of the roughness; I assumed that snow melts latest in valleys and earliest on ridges because of the differences in micro-climates. (8) Lake; This mostly methodical parameter might not strongly affect the snow coverage, otherwise it could hint at an influence of either the magnitude of the wiggling of the camera or a change in the microclimate because of the volume of the nearest water body. (9) Analysis quality; This variable described if the analysed snow coverage fitted the visually expected measures. I expected a difference in their predictive power. (10) Number of red pixels; Nest sites, of which only a small part was visible, could influence the analysed measure. (11) Number of analysed days; This assumedly affected the computed statistical measures as a snowfall event could only be detected if it was recorded. (12) Observation year; I expected a large between-year

variation as the weather differed.

To fit this small data set with numerous variables more easily, I transformed all variables using the function *stats::poly* with orthogonal polynomials of the first and second degree. I chose a linear mixed effect model because of the random effect of the year[6] and the assumed normal distribution of the observations [F. Korner-Nievergelt, Roth, et al. 2015].

Model

I created the following linear mixed effect model (LMEM) in R to investigate the influence of above factors on the first snow-free day, last snow-covered day and number of days temporary snow-free:

$$
\begin{aligned}
lmer(snow\text{-}factor \sim\ & NDVI.median.l + NDVI.median.q+ \\
& NDVI.perc.l + NDVI.perc.q+ \\
& elevation.median.l + elevation.median.q+ \\
& aspect.median.l + aspect.median.q+ \\
& aspect.perc.l+ \\
& slope.median.l + slope.median.q+ \\
& roughness.median.l + roughness.median.q+ \\
& lake + analysis\,quality+ \\
& number\,of\,red\,pixels + number\,of\,analysed\,days+ \\
& (1|Observation\,year),\ data = dat,\ REML = T)
\end{aligned}
\tag{11}
$$

Let snow-factor be (1) first snow-free day, (2) last snow-covered day or (3) number of days which are temporary snow-free. If the variable name contained two dots, it was composed as follows: factor.measure.degree; Where the factor was the analysed environmental factor (e.g. slope), the measure was either the median or the 90% percentile distance (perc) and the degree was the orthogonal polynomial (l for linear or q for quadratic). This nomenclatorial rule applies to all models.

---

[6] Data collected in a second year does not add as much information to the model as the first year did [Meier 2018].

2.5.2 Influence of the environmental factors on the likelihood of a presence point

I created the following three models to investigate the influence of the described environmental factors, the snow coverage and the between-year variation on the nest site choice of the northern wheatear.

Maximal generalized linear mixed model (GLMM) - without snow
The following maximal GLMM described the influence of the observation year on the environmental factors.

Data
I analysed 468 data points which were located outside of the Lago Ritom, see Table 5.

	2013	2014	2015	2016	2017	2018
Absence points	17	44	33	49	42	56
Presence points	16	40	32	47	38	54

Table 5: A summary of the analysed number of presence and absence points per year.

I wished to report the effect of the maximal model [F. Korner-Nievergelt and P. Korner-Nievergelt 2019]. Therefore, I fitted eleven computed environmental variables, which were not strongly correlated to the other variables [Meier 2018]: (1) Median of the aspect. (2) 90% percentile distance of the aspect; I excluded the 90% percentile distance of the roughness from analysis because of its high correlation to the 90% percentile distance of the aspect, see Figure 66. (3) Median of the elevation. (4) Median of the NDVI. (5) 90% percentile distance of the NDVI. (6) Median of the roughness. (7) Median of the slope; I excluded the 90% percentile distance of the elevation from analysis because of it highly correlates to the median of the slope, see Figure 66. (8) 90% percentile distance of the slope. (9) Cumulated distance to the three nearest nests. (10) Distance to the nearest path. (11) Observation year.
To fit this data set with its numerous variables more easily, I transformed all variables using the function *stats::poly* with orthogonal polynomials of the first and second degree. For the expected effects on the prediction of the variables of the model, see 2.3.3 to 2.3.9.

Model
I required a model to not only predict a binomial distribution, i.e. to differentiate between presence and absence point, but also to include a random effect of the observation year. Therefore, I chose the

following GLMM.

$$
\begin{aligned}
glmer(Occupation \sim \ & aspect.median.l + aspect.median.q+ \\
& aspect.perc.l+ \\
& elevation.median.l + elevation.median.q+ \\
& NDVI.median.l + NDVI.median.q+ \\
& NDVI.perc.l + NDVI.perc.q+ \\
& roughness.median.l + roughness.median.q+ \\
& slope.median.l + slope.median.q+ \\
& slope.perc.l+ \\
& cumulated\ distance\ to\ the\ three\ nearest\ nests+ \\
& distance\ to\ the\ nearest\ path+ \\
& (1|Observation\ year), data = dat,\ family = binomial)
\end{aligned}
\tag{12}
$$

Model assumptions

I checked the following plots to test if the model assumptions were met. I checked the residual plots for shrinkage (Tukey-Anscombe plot), tested the normal distribution of the residuals (Normal Quantile-Quantile plot), described the residual variance (Scale-Location) and visualised the influence of data points on the outcome (Leverage) [F. Korner-Nievergelt, Roth, et al. 2015]. The distribution of the residuals of a binomial model may deviate from the diagonal line [F. Korner-Nievergelt and P. Korner-Nievergelt 2019]. To determine if the present deviation is tolerable, a comparison of the found distribution to eight randomly drawn normal distributions was recommended [F. Korner-Nievergelt, Roth, et al. 2015]. The presence of an unexplained spatial autocorrelation was tested by comparing the found spatial distribution of residuals to the spatial distribution of eight modelled residual maps [F. Korner-Nievergelt and P. Korner-Nievergelt 2019]. Furthermore, the Bayesian p-value checked for more extreme values in the simulated data than in the observed data [F. Korner-Nievergelt, Roth, et al. 2015]. A Bayesian p-value of 0.5 indicates a perfect model fit as "the test statistics from the observed data is well in the middle of the ones from the replicated data" [F. Korner-Nievergelt, Roth, et al. 2015]. I computed the Bayesian p-value for the proportion of absence points, the 10% quantile of residuals and the 90% quantile of residuals. Finally, I checked for underdispersion and overdispersion by comparing the residual deviance and the degrees of freedom [F. Korner-Nievergelt, Roth, et al. 2015].

Simulation

I drew 10000 random samples of the posterior distribution from the model parameters using equation 13.

$$
bsim = sim(mod, n.sim = 10000)
\tag{13}
$$

Then, the 2.5%, 50% and 97.5% quantiles were calculated from these simulated values. Between the 2.5% and 97.5% quantiles spanned the 95% credible interval [F. Korner-Nievergelt, Roth, et al. 2015]. Furthermore, the 2.5% and 97.5% quantiles of a factor indicated evidence for a strong effect if they shared their algebraic signs, i.e. both were positive or negative.

Maximal generalized linear model (GLM) - without snow

I created the following maximal GLM without snow to analyse which environmental factors affected the nest site choice of the northern wheatear.

Data

I analysed the same data points as in the GLMM, see Table 5.

To fit the above described data set with its eleven variables more easily, I transformed all variables using the function *stats::poly* with orthogonal polynomials of the first and second degree.

Model

This model needed to predict a binomial distribution while there was no requirement of including the observation year. Thus, a GLM predicting the occupation of a point was sufficient.

$$
\begin{aligned}
glm(Occupation \sim\ & aspect.median.l + aspect.median.q+\\
& aspect.perc.l+\\
& elevation.median.l + elevation.median.q+\\
& NDVI.median.l + NDVI.median.q+\\
& NDVI.perc.l + NDVI.perc.q+\\
& roughness.median.l + roughness.median.q+\\
& slope.median.l + slope.median.q+\\
& slope.perc.l+\\
& cumulated\ distance\ to\ the\ three\ nearest\ nests+\\
& distance\ to\ the\ nearest\ path+\\
& data = dat, family = binomial)
\end{aligned}
\tag{14}
$$

Model assumptions

I tested the same model assumptions as above.

Simulation

I drew 10000 random samples of the posterior distribution from the model parameters using equation 13 and calculated the 2.5%, 50% and 97.5% quantiles, as above.

Maximal generalized linear model (GLM) - with snow

Originally, I wished to determine the effect of all created environmental factors including the snow coverage on the nest site choice of the northern wheatear under the consideration of the between-year variation of these factors. Unfortunately, I encountered difficulties during the model convergence. Therefore, I decided to reduce this model so that it only described the effect of some environmental factors without taking the between-year variation into account [Meier 2018].

Data

I analysed the same data points as in Table 4: the GLM with snow was run on 69 nests for which I could calculate the snow cover. I only fit the variables (1) median of the aspect, (3) median of the elevation, (4) median of the NDVI, (6) median of the roughness, (7) median of the slope, (9) cumulated distance to the three nearest nests, the interaction of the median of elevation with the median of slope and the two snow coverage variables first snow-free day and last snow-covered day. To do so more easily, I transformed all variables using the function *stats::poly* with orthogonal polynomials of the first and second degree. Additionally, I log-transformed the cumulated distance to the three nearest nests.

Model

I required a model to predict a binomial distribution and thus chose the following GLM.

$$
\begin{aligned}
glm(Occupation \sim aspect_median.l+ \\
elevation.median.l + elevation.median.q+ \\
NDVI.median.l+ \\
slope.median.l + slope.median.q+ \\
roughness.median.l+ \\
cumulated\ distance\ to\ the\ three\ nearest\ nests+ \qquad (15) \\
first\ day\ snow\text{-}free+ \\
last\ day\ snow\text{-}covered+ \\
elevation.median.l : slope.median.l, \\
data = dat, family = binomial)
\end{aligned}
$$

Model assumptions

I tested the same model assumptions as above.

Simulation

The 10000 random samples drawn of the posterior distribution were used to compute the 2.5%, 50% and 97.5% quantiles, as above.

3 RESULTS

3.1 The data

The main purpose of the collected data was to model the nest site choice of the northern wheatear, see 2.5.2. Nonetheless, the collected data reported the similarity of the NDVI compared to the sampled land cover, the distance to the nearest nest and the nest density.

3.1.1 Normalized difference vegetation index (NDVI)

The comparison of the sampled land cover and different NDVI median classes disclosed a positive correlation for vegetated ground and a negative correlation for unvegetated ground. The vegetated ground equalled vegetation <5cm + vegetation 5 - 15cm + vegetation >15cm + shrubs. The unvegetated ground summed bare ground and rocks.

cor.test(NDVI, x)	cor	t	df	95% confidence interval	p-value
Vegetation <5cm	+0.28	+4.34	230	[+0.15, +0.39]	<0.01
Vegetation 5-15cm	+0.20	+3.12	230	[+0.07, +0.32]	<0.01
Vegetation >15cm	+0.19	+2.98	230	[+0.07, +0.31]	<0.01
Shrubs	+0.08	+1.16	230	[- 0.05, +0.20]	0.25
Bare ground	- 0.58	- 10.87	230	[- 0.66, - 0.49]	<0.01
Rocks	- 0.39	- 6.43	230	[- 0.49, - 0.28]	<0.01
Vegetated ground	+0.54	+9.79	230	[+0.44, +0.63]	<0.01
Unvegetated ground	- 0.61	- 11.54	230	[- 0.68, - 0.52]	<0.01

Table 6: The NDVI correlated strongest to the sampled land cover classes when they were summarised to vegetated and unvegetated area. The NDVI correlated positive to vegetated ground and negative to unvegetated ground. The NDVI correlated slightly positive to the land covers vegetation <5cm, vegetation 5-15cm and vegetation >15cm. Shrubs displayed no correlation. A negative correlation existed between NDVI and bare ground or rocks.

These trends of correlation, or absence thereof, to the individual land cover classes are partly visible in Figure 33. The sampled land cover equalled the ratio of that particular land cover class compared to the total area of the nest site. The NDVI median describes the median of all NDVI values occurring in the nest site.
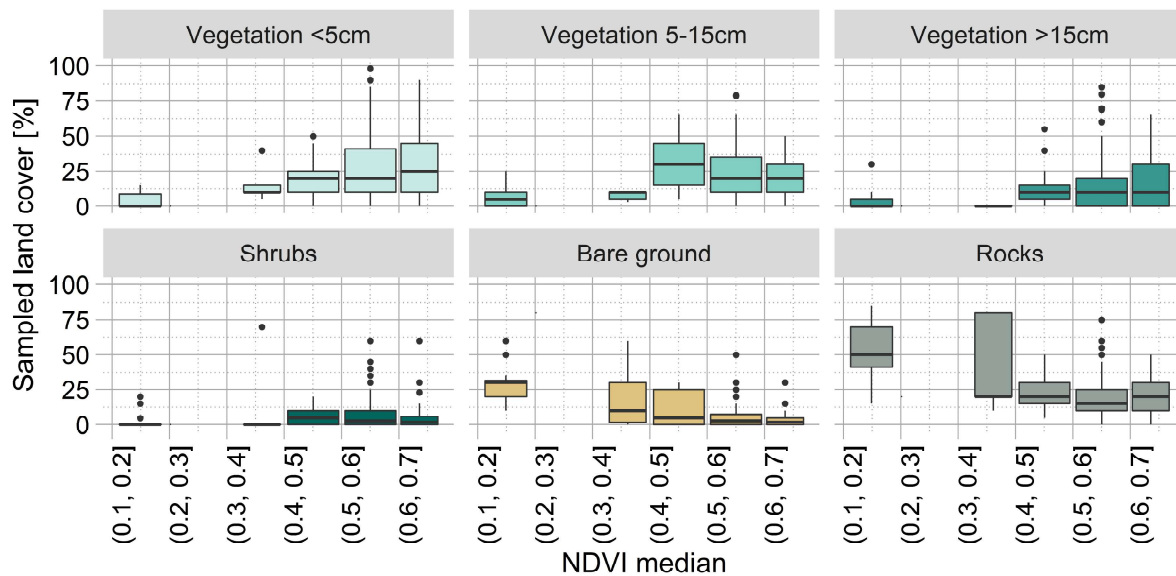
Figure 33: The proportion of land covers vegetation <5cm, vegetation 5-15cm and vegetation >15cm increased slightly with an increasing median of NDVI. The ratio of bare ground and rocks decreased with increasing median of NDVI.

3.1.2 Distance to the nearest nest

In Val Piora, the distance to nearest two nests (slant range) ranges from 15m to 36m, see Table 3. The average minimal distance to the nearest nest in the Val Piora is 26m.

3.1.3 Nest densities

Under the assumption of uniformity, a density of ~4 nests per 10ha exists in Val Piora.
The computed local densities varied between 1 and 7 nests per 10ha per year and the calculated overall densities ranged from 0.2 to 0.5 nests per 10ha per year, see Table 7.

Year	Local minimum	Local mean	Local maximum	Overall
2013	1	1.7	4	0.2
2014	1	2.5	5	0.4
2015	1	2.4	6	0.3
2016	1	2.8	7	0.5
2017	1	2.3	4	0.4
2018	1	2.5	5	0.5
Average	1	2.4	5.1	0.4

Table 7: The densities were computed in nests per 10ha. The mean local nest densities per year range between 1.7 and 2.8. The maximal found density was 7 nests per 10ha in 2016, which was a year with the highest overall density.

3.2 The snow coverage algorithm

The algorithm created two types of results: Firstly, as proof of principle, it was possible to extract information of the snow coverage course at the nest site by overlaying the *1-Model* and the *2-Photo*, see 3.2.1. Further, the pixels were matched between the *1-Model* and the *2-Photo*, see 3.2.2. Secondly, this extracted information of snow coverage could be used in my statistical models, see 3.3.1 and 3.3.2.

3.2.1 Accuracy of overlay

Because of the resemblance of the *1-Model* and the reference *2-Photo*, it was in all cases possible to overlay them. However, the manual choice of the GCPs introduced inconsistencies and inaccuracies. To quantify the performance of this function, I overlaid 15 times the same *2-Photo* and the same *1-Model* of the Lago Cadagno recorded in 2017 and chose the same four GCPs. The overlay function computed a scaling factor and an offset in x- and y-direction. The repetitions of 15 perfect fits produced assumedly 15 times the same values. Table 8 shows the summary of these repeatedly computed scaling factors and offsets.

	scale_x	scale_y	offset_x	offset_y
Min	0.95	0.87	-66.24	-188.44
Max	1.00	0.88	-20.67	-177.83
Mean	0.97	0.87	-41.44	-181.60

Table 8: The scaling factor and offset varied stronger in x-direction than in y-direction. The scaling factor is without unit.

The offset was given in pixels, however the unit metre was more intuitively interpretable. Since the covered ground per pixel varied with the distance of the photographed ground to the camera, there is no direct transformation. To approximate the dimensions of one pixel, I sampled the size of the nest site on six *1-Models*. The diameter of a nest site was always 50m. The mean size of a pixel of these six *1-Models* was 1.4m in x-direction and 2.4m in y-direction. Thus, a difference of 45.57 pixels in x-direction equalled 63.6m and a difference of 10.61 pixels in y-direction equalled 25.5m.

3.2.2 Accuracy of snow coverage per nest site and image

The comparison of the snow coverage of the nest site recorded by two cameras disclosed a similar snow coverage course, with less noise recorded at the Lago Ritom. First, the dates of distinct events, such as fresh snow, differed between the two analysis. I estimated the temporal accuracy by comparing the date of the first snow-free day, respectively the last snow-covered day between the two recordings. The found accuracy of these specific nest sites was one week. Second, the results of the two analysis were not perfectly comparable. Once the nest was placed to the left of the visually determined nest location and once to the right. Thus, the two analysis compared not the exact same location.

3.2.3 Accuracy of snow coverage per nest site and season

The results of the statistical models proofed the principle for the variables first snow-free day, last snow-covered day and number of temporary snow-free days, see 3.3.1. I qualitatively analysed various plots such as Figure 29. The three main findings are: (1) The depicted snow coverage usually increased after an anticipated snow event, e.g. mid-April or end of April in Figure 29. Thus, the late snow events which occurred across the entire study area were recorded by the cameras. (2) There was a high uncertainty in snow coverage depending on the illumination condition of the *2-Photos* and thus they spanned over more than half of the possible range, e.g. end of March or April. (3) The median hardly ever laid in the visual middle of the 10% percentile and the 90% percentile.

3.3 The statistical models

3.3.1 Plausibility of the computed snow coverage variable

This proof of principle determined observation year, median of slope, median of elevation and median of NDVI as the most influential factors on the computed snow coverage variables. These strong effects of environmental factors might indicate that the method described the snow coverage. A strong influence of methodological factors, such as the recording duration, would hint at a suboptimal algorithm. I simulated, using *arm::sim*, 2000 random samples of the posterior distribution of the three linear mixed effect models (LMEM).

Each of the three discussed LMEM modelled either first snow-free day, last snow-covered day or number of days temporary snow-free. All models investigated the effect of the variables median of NDVI, 90% percentile distance of NDVI, median of elevation, median of aspect, 90% percentile distance of aspect, median slope, median of roughness, lake (camera location), analysis quality, number of red pixels (size of analysed area), number of analysed days and observation year. The detailed effects of the analysed variables on the snow coverage variables are listed in 6.3.5. For better comparison, I numbered the found effects according to the expectations in 2.5.2.

First snow-free day
The main points of the modelled first snow-free day are shortly described below. An strong between-year variation of the first snow-free day was indicated. A nest site was earlier snow-free (small value for first snow-free day) if the area was steep (large linear term of the median of the slope).
A strong negative linear effect was indicated for (6) the median of slope.
An absence of strong negative linear effect was expressed for (3) the median of elevation, (7) the median of roughness, (9) analysis quality and (10) number of red pixels. Evidence for an absence of a strong positive linear effect was found for (1) the median of NDVI, (2) the 90% percentile distance of NDVI and (11) the number of analysed days. It further suggested an absence of a strong quadratic effect for (1) the median of NDVI, (3) the median of elevation, (4) the median of aspect and (7) the median of roughness. An absence of linear trend was suggested for (4) the median of aspect, (5) the 90% percentile distance of aspect, (8) Lago Ritom and (8) Lago di Tom. The model showed an absence of quadratic trend for (2) the 90% percentile distance of NDVI and (6) the median of slope.

Last snow-covered day
The main points of the modelled last snow-covered day are shortly described. A strong between-year variation of the last snow-covered day was indicated. A nest site was earlier permanently snow-free on lower median elevations.
A strong positive linear effect was indicated for (3) the median of elevation.
The computed LMEM found evidence for an absence of a strong negative quadratic effect was shown for (3) the median of elevation. No strong negative linear effect was found for (7) the roughness, (8) Lago di Tom and (11) number of analysed days. Furthermore, the model revealed evidence for an absence of a strong positive linear effect for (4) the median of aspect and (6) the median of slope. No strong positive quadratic effect was indicated for (2) the 90% percentile distance of NDVI, (6) the median of slope and (7) the median of roughness. An absence of linear trend was found for (1) the median of NDVI, (2) the 90% percentile distance of NDVI, (5) the 90% percentile distance of aspect, (8) Lago Ritom, (9) analysis quality and (10) number of red pixels. Evidence for an absence of quadratic trend was indicated for (1) the median of NDVI and (4) the median of aspect.

Number of days temporary snow-free
The main points of the modelled number of days temporary snow-free are shortly described. A existing between-year variation was indicated. The first snow-free day and last snow-covered day were closer together if the median of slope was steep and and the median of NDVI high.

A strong positive linear effect was expressed for (6) the median of slope and a strong positive quadratic effect for (1) the median of NDVI. Evidence for the absence of a strong positive linear effect was indicated for (1) the median of NDVI, (9) analysis quality and (10) number of red pixels. Additionally, evidence for the absence of a strong negative linear effect was identified for (8) Lago Ritom, (8) Lago di Tom and (11) number of analysed days. No strong negative quadratic effect was found for (3) the median of elevation, (4) the median of aspect and (7) the median of roughness. The computed LMEM described the absence of a linear effect for (2) the 90% percentile distance of NDVI, (3) the median of elevation, (4) the median of aspect, (5) the 90% percentile distance of aspect and (7) the median of roughness. No quadratic effect was indicated for (2) the 90% percentile distance of NDVI and (6) the median of slope.

3.3.2 Influence of the environmental factors on the nest site choice

Maximal generalized linear mixed model (GLMM) - without snow
The GLMM, formula 12, scrutinized the between-year variation of the analysed environmental factors, see 6.3.6. The random factor observation year disclosed an absence of a between-year variation for the scrutinized environmental factors. The original model output indicated effects of the variables median of the aspect, 90% percentile distance of the aspect, median of the elevation, median of the NDVI, 90% percentile distance of the NDVI, median of the roughness, median of the slope, 90% percentile distance of the slope, cumulated distance to the three nearest nests and distance to the nearest path. The effects of these variables were scrutinised with the following GLM without considering the observation year:

Maximal generalized linear model (GLM) - without snow
This GLM, formula 15, analysed the effect of the median of aspect, the 90% percentile distance of aspect, the median of elevation, the median of NDVI, the 90% percentile distance of NDVI, the median of roughness, the median of slope, the 90% percentile distance of slope, the cumulated distance to the three nearest nests and the distance to the nearest path. The detailed effects of the investigated variables on the nest site choice of the northern wheatear in the Val Piora are listed in 6.3.7. The most important points of the original model output are briefly summarised.
The model assumptions were met [F. Korner-Nievergelt, Roth, et al. 2015]. The residual plots revealed no violation of the model assumptions (see Figure 82). The Q-Q plot of the model was in line with eight simulations (see Figure 84). The found spatial autocorrelation was recreated by including the distance to the nearest path into the model [F. Korner-Nievergelt, Roth, et al. 2015]. The observed residuals could not be discriminated from the residuals of eight simulations (see Figure 83). The computed Bayesian p-value indicated suitable representations of the proportion of absence points, of the 10% quantile of residuals and of the 90% quantile of residuals (see 6.3.7). Only a slight underdispersion was detected as the residual deviance was smaller than the degrees of freedom, i.e. data varies less than expected [F. Korner-Nievergelt, Roth, et al. 2015].
Strong negative linear effects were found for the median of elevation, the median of NDVI, the median of slope, the 90% percentile distance of slope, the cumulated distance to the three nearest nests and the distance to the nearest path. The model indicated strong quadratic effects for the median of aspect, the 90% percentile distance of NDVI and the median slope. A strong positive linear effect emerged for the median of roughness. Evidence for an absence of a strong positive linear effect existed for the median of aspect and the 90% percentile distance of aspect. The model expressed an absence of a strong positive quadratic effect for the median of elevation and the median of roughness. Evidence for an absence of a strong negative linear effect surfaced from the 90% percentile distance of NDVI and a strong negative quadratic effect from the median of NDVI.
The fitted model predicts a maximum for S- to SW-exposed slopes, elevation between 1900 and 2100m a.s.l., median of NDVI values of -0.1 to +0.2, NDVI ranges of 0.4 to 0.6, roughness of 0.03 to 0.08, 15° to 30° steep slopes, little variation in the incline of these slopes, sites close to the three nearest nests and sites close to the nearest path, see Figures 34 to 42. These computed strong effects were visualised in the following effect plots.
An effect plot visually describes the effects of one variable on the nest site choice if all other variables are constant [F. Korner-Nievergelt and P. Korner-Nievergelt 2019]. The simulated values (x-axis) of this variable span over the range of the original variable. The index of success (y-axis) indicates the likelihood of a presence point.[7] The value of a sampled presence point (Index of success = 1) or absence point

---

[7] If the used absence points were true-absence points instead of pseudo-absence points, the model would have predicted the likelihood

(Index of success = 0) is displayed as jittered black circles. The black line displays the fitted (simulated) values. The grey envelope spans between the 2.5% percentile and 97.5% percentile of all simulated models; a narrow envelope indicate a confident model.
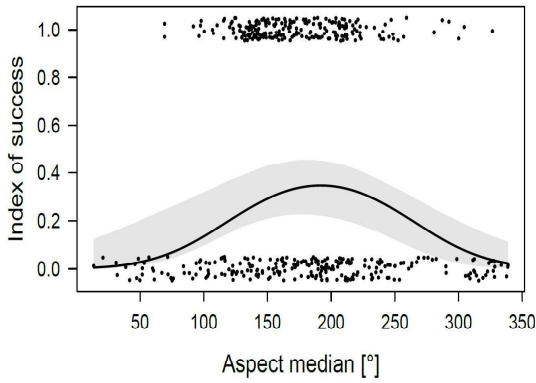

Figure 34: The fitted maximum of linear effect of the median of aspect spans between S (180°) and SW (225°). The maximum of the envelope stretches between SE (135°) and SW. The insecurity of the prediction is uniform. The presence points scatter around the eastern end of the predicted maximum.
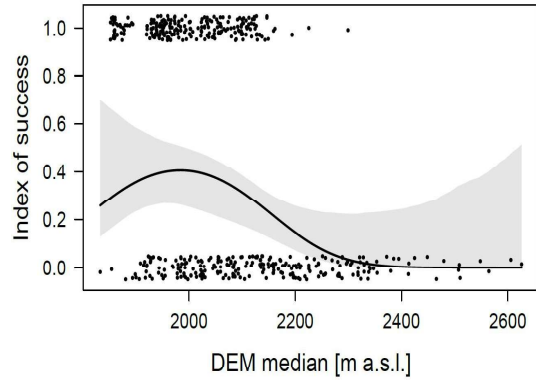
Figure 35: The fitted maximum of the median of elevation spans between 1900m a.s.l. and 2100m a.s.l. The uncertainty of the prediction increases towards the either end of the possible elevation values.
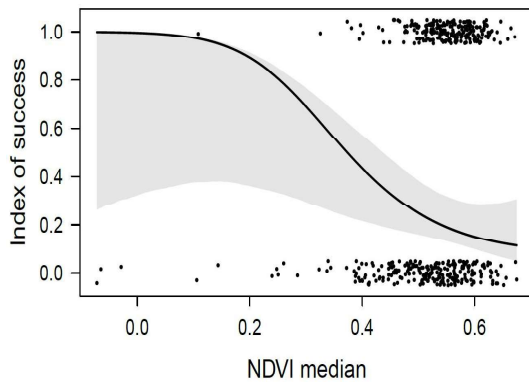
Figure 36: The fitted maximum of the linear effect of the median of NDVI stretches between -0.1 and +0.2. The uncertainty largely increases towards the lower end of the possible values. The presence and absence points concentrate around the predicted minimum.

Figure 37: The fitted maximum of the 90% percentile distance of NDVI and the maximum of the envelope span between 0.4 and 0.6. The uncertainty of the prediction increases towards the top end of the range. The of the presence and absence points focus towards the predicted (left) minimum of the range.

of a presence point itself.
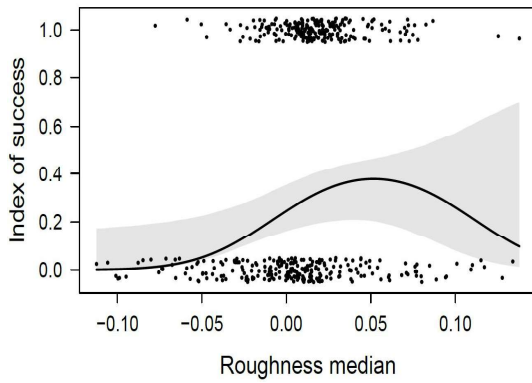
Figure 38: The fitted maximum of the median of roughness ranges between 0.03 and 0.08. The uncertainty of the prediction increases towards the top end of the occurring values. The absence and the presence points cluster to the left of the predicted maximum.
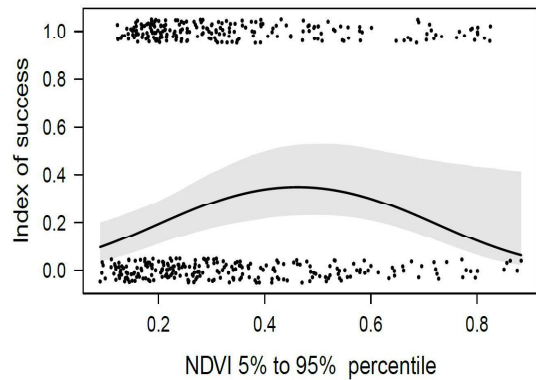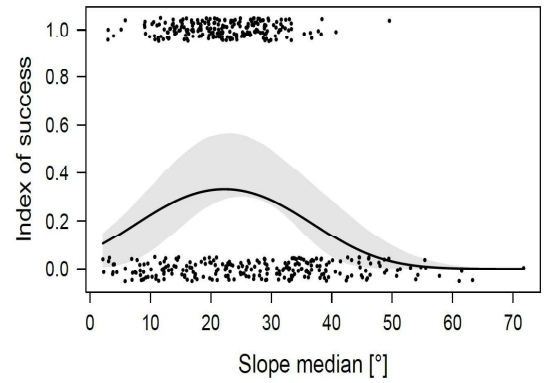

Figure 39: The fitted maximum of the median of slope and the maximum of the envelope span between 15° and 30°. The otherwise uniform uncertainty of the prediction disappears above 50°.
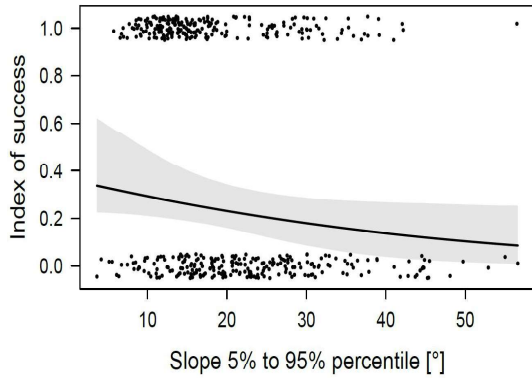
Figure 40: The fitted values of the 90% percentile distance of slope and the envelope display a slightly decreasing index of success. The uncertainty of the envelope slightly increases towards the lower end of the range.
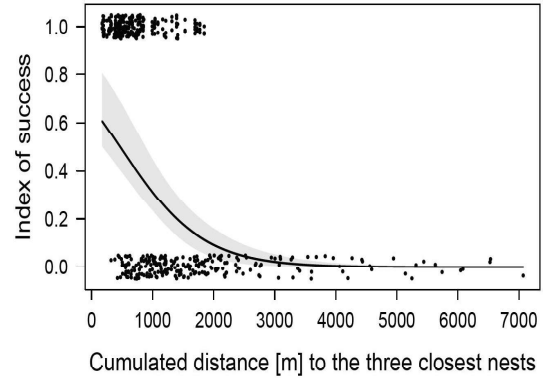
Figure 41: The fitted values of the cumulated distance to the three nearest nests reveal a strongly decreasing index of success. The uncertainty and possibility of success disappears for distances >3000m.
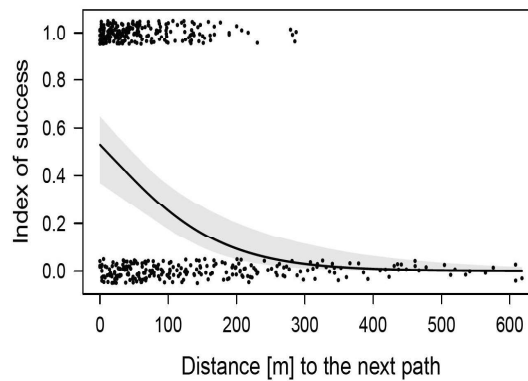
Figure 42: The index of success of the fitted values of the distance to the nearest path and of the envelope decreases. The uncertainty disappears for distances >400m.

Maximal generalized linear model (GLM) - with snow

This GLM, formula 15, analysed the effect of some environmental factors and the effect of the snow coverage variables first snow-free day and last snow-covered day. The most important points are shortly summarised. The present model was the maximal model still meeting the model assumptions, see Figure 85. The cook distance never exceeded one, the Q-Q-plot was similar to the eight random models and so was the spatial autocorrelation. The Bayesian p-value indicated a good model fit. Underdispersion was detected, i.e. the data varies less than expected [F. Korner-Nievergelt, Roth, et al. 2015].

Strong negative linear and quadratic effects were found for the median of slope. Evidence for the absence of a strong negative linear effect exists for the median of aspect, the interaction of the median of elevation with the median of slope, the median of NDVI and the last snow-covered day. Evidence for the absence of a strong positive linear effect emerged for the variables median of elevation, median of roughness, cumulated distance to the three nearest nests and the first snow-free day. No strong positive quadratic effect was found for the median of elevation.

The fitted model predicts a strong maximum only for slopes between 25° and 30°, see effect plot in Figure 45. The additional snow coverage variables display no strong effect which is displayed in the effect plots in Figure 43 and 44.
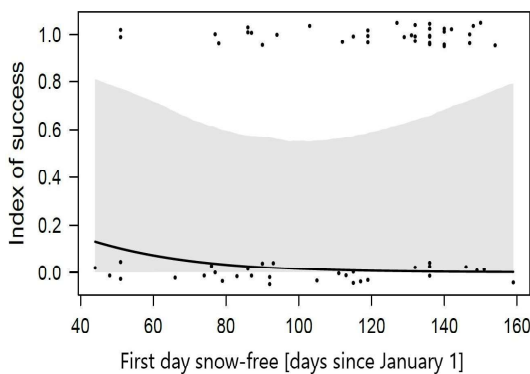
Figure 43: The fitted first snow-free day slightly decreases. The insecurity of the prediction is enormous. The scarce presence points scatter around the top end of the possible values.
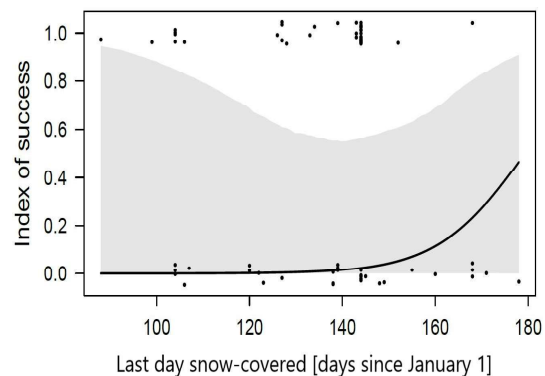
Figure 44: The fitted last snow-covered day strongly increases for values >150 (~end of May). The insecurity of the prediction is immense. The scarce presence points focus in the centre of the entire range.
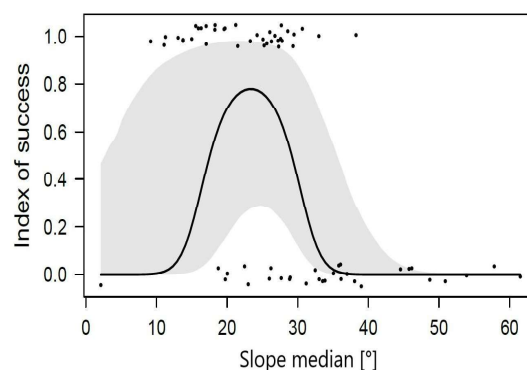

Figure 45: The fitted maximum of slope spans between 25° and 30°. The maximum of the envelope ranges between 10° and 30°. The insecurity of the prediction is large for values <30°.

4 DISCUSSION

4.1 The data

The data for the normalized difference vegetation index (NDVI) was sampled at the end of August. Consequently, it described a different stage of vegetation than what the northern wheatears encountered at their arrival in spring. Thus, the here created NDVI possibly describes rather the effect of the NDVI on the nest site selection of the subsequent season [Arlt and Pärt 2008a]. On the other hand, bare ground in August was assumedly also bare ground in spring. Thus, this NDVI may still contribute to the approximation of what the wheatears encountered in spring. Furthermore, such a late sampling date of the NDVI allowed the recorded vegetation to grow all summer. Therefore, I expect a less pronounced difference between densely and sparsely vegetated areas in the recorded data than if the same areas were sampled in spring, see Figure 50. Originally, I created the NDVI to approximate the recorded land cover data. However, the NDVI seemed unable to differentiate the sampled six classes, see 3.1.1. Still, it correlated to the ratio of vegetated respectively unvegetated ground. Thus, the here presented NDVI added details to the remotely described landscape despite the suboptimal recording date.

The small cumulated distance to the three nearest nests (CD3N) hints at the spatial components of the social behaviour of the northern wheatear. This measure is possibly influenced by the number of data points per year. Only half as many nests were found in 2013 as in any other season. Consequently, the CD3Ns in 2013 were larger than in the other years, see Figure 15. The maximal local nest density was larger than the expected nest density of the uniform distribution of the average CD3N, see 3.1.3. This suggested clustered or randomly distributed nests. The visual impression indicates clustering. Furthermore, the reported nest densities ranged between $2.5/km^2$ and 5/10ha in the Alps [Catzeflis 1976; Glutz von Blotzheim 1962]. The found maximal local nest densities exceeded the maximal literature value in two years and the mean local nest density ranged between 1.7 and 2.8 nests per 10ha per year. Therefore, I conclude that the northern wheatears in the Val Piora prefer to nest in small clusters, as Arlt and Part 2007 suggested. Maybe, breeding northern wheatears prefer clustered nests as the patches of suitable habitats can accommodate multiple wheatear pairs (local ideal free distribution [Fretwell and Lucas 1969]) or because they defend their breeding grounds together against predators.

The two nearest nests were found in 2013, a year where the Swiss northern wheatear population was low [Schweizerische Vogelwarte Sempach n.d.]. The distance of 15m between two nests considerably undercut the previously reported nest proximities of barely 100m in the Swiss Alps [Glutz von Blotzheim 1962]. On average, there was a distance of 26m between the nearest two nests per year. Thus, the choice of 25m as radius around a nest site seems appropriate to investigate environmental factors and the snow coverage. Because of these low values, the Val Piora appears densely populated by northern wheatears. In accordance with the literature, the found average distance was 178m between two nests [Corti and Melcher 1958; Glutz von Blotzheim 1962].

Contradicting my assumption, the northern wheatears preferred path proximity, see Figure 16b. However, the northern wheatears were described as shy [Maumary et al. 2007]. Therefore, this surprising result might be affected by a sampling bias in geographic space [Graham et al. 2004]. Consequently, the likelihood of detection might be lower than expected, especially far from the paths. Alternatively, the northern wheatears in the Val Piora could be less shy as they could be accustomed to the constant presence of hikers.

4.2 The snow coverage algorithm

The snow coverage algorithm successfully measured the snow coverage on a nest site and found a link to the measured weather data. The underlying setup was chosen because of its relatively simple equipment, large spatial range and low price. The **preparation** of the data (*1-Model*, *2-Photo* and *3-Suitability*) influenced the further methods. The precision of the snow coverage algorithm was affected by the match between the 3D-visualised DEM and the *2-Photo*. A major challenge was that the coordinates did not allow to precisely define the right point of view on the DEM. Amongst others, there often was an imprecision in altitude, which needed manual adjustment. The detail of the *1-Model* approximated the detail of the camera. Unfortunately, complete congruence was not achieved.

In general, the cameras were intended to record the same detail throughout the season. In practice, the displayed scenes varied slightly over time, see Figure 46. These differences in detail were especially pronounced in the camera recording Lago Cadagno. The cameras were mounted on trees, which moved presumedly in the wind and under the snow load. An additional movement of the camera in its holder can not be excluded.

(a) At the start of the footage, the camera recorded this scene. Later in season, it recorded Figure 46b and afterwards Figure 46c.

(b) The horizontal offset compared to Figure 46a can be seen in the smaller space between the left margin and the leftmost larch or the larger space between the right margin and the flank of Pizzo Taneda.

(c) The vertical offset compared to Figure 46a is visible in the smaller space between the top margin and the horizon.

Figure 46: These three images were taken at Lago Cadagno in 2016. They exemplify the spatial inaccuracy of wiggling cameras.

Because of these wiggling cameras, the index of the nest site (red-pixel-index) matched varying *2-Photo* details. Thus, the spatial inaccuracy was enormous but not estimated. I conclude firstly, the wiggling cameras introduced random noise and secondly, the choice of a reference *2-Photo* (for rotation or the overlay function) influenced the outcome of the snow coverage analysis.

The chosen camera type seemed challenged by the harsh weather conditions of the alpine environment. The light-weighted cameras were easily installed, recorded almost the entire season and required little maintenance. However, I noticed the following three challenges. Firstly, some recording times were skipped, especially in cold weather. Furthermore, the camera recorded *2-Photos* as bright as day after midnight in summer. Thus, the clock of the camera appeared to propagate the temporal errors it endured in winter. Consequently, the timestamp of the camera is merely a temporal guide value. Secondly, most *2-Photos* recoded on a sunny day were overexposed. Especially overexposed were white winter scenes under a blue sky. On such scenes, the greyish-red rocks were indistinguishable from snow or gravel beaches. The surrounding snow overshone even previously snow-free areas. Thirdly, a dark halo appeared around the centre of a few *2-Photos*; Pixels near the margins displayed lower RGB values than pixels near the centre for the same feature. Consequently, marginal pixels of snow were

overlooked while the central pixels were correctly classified as snow. Thus, these cameras were not the most adapted to record the course of the snow coverage.

The creation of *3-Suitability* or the RGB threshold of snow underlied the assumption, that one RGB threshold would suffice to distinguish snow from other surfaces under all illumination conditions of fine weather. The illumination of the *2-Photos* suitable for analysis (recorded in fine weather) varied considerably, see Figure 47.
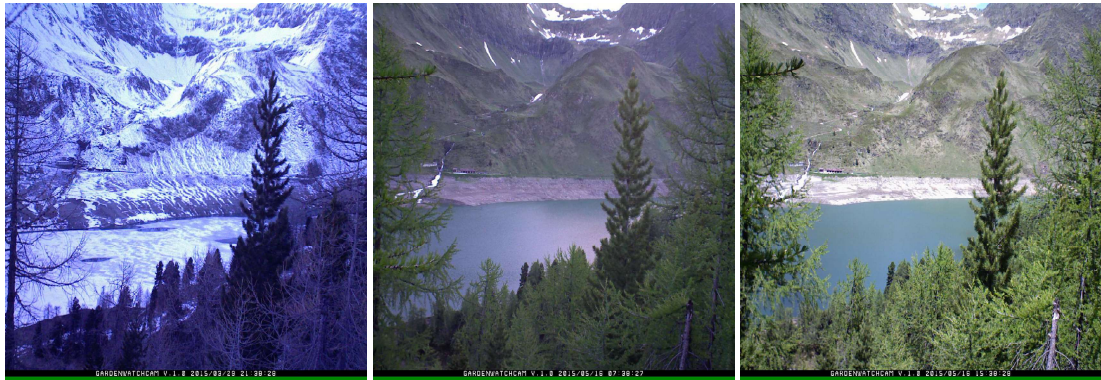
(a) This image is bluish-shadowy. Such an extreme illumination condition originates probably in the low amount of daylight. Thus, snow was not white and would not always be detected with the chosen threshold.

(b) This image depicts neutral illumination condition. Snow should not only be well detectable but also well distinguishable from other surfaces. Only the white water would be mistakenly labelled snow.

(c) This image was recorded on a bright sunny day. Concrete surfaces, hiking trails or the beach appear as bright as snow and are not distinguishable.

Figure 47: These three images were labelled suitable for analysis as the weather was fine and no shadows obscured parts of the opposite shore. They depict the range of possible illumination conditions.

Figure 47 discloses the diversity of the spectral appearance of snow and other surfaces under varying illuminations on the footage. The creation of a single threshold which detected snow in all cases and simultaneously distinguished snow from all other surfaces was impossible. I expected that concrete, gravel and water sometimes were labelled snow (false positive), whereas snow was occasionally overlooked when snow was in the shadow, dirty or the illumination appeared shadowy-bluish (false negative). Nonetheless, I tried to create a RGB threshold, which took minor special cases of illumination into account. The strongest special cases should be removed in snow coverage per nest site and season with a slightly smoothing statistical measure. However, the multitude of such extreme illumination conditions was underestimated. Thus the initially chosen combination of RGB threshold, selection of suitable *2-Photos* and statistical measure resulted in many influential outliers. Consequently, adaptions were needed and described below.

One of the last manual steps of the algorithm was the selection of at least three ground control points (GCPs) on the recorded scene in the **overlay** function. This method was simple in execution and tailored to the existing data. Additionally, it allowed flexibility in the choice of area of interest within the recorded scene, e.g. a cluster of nests. Instead of a mediocre overall accuracy, a high accuracy at these nests could be chosen at the price of a worse overall accuracy. On the other hand, the selection of these GCPs was restricted by the depicted features on the *2-Photo* and *1-Model*. Only well-chosen GCPs fit the area of interest visually satisfying. The selection of GCPs involved the trial-and-error principle.

The manual choice of the same GCPs affected the computed scaling factor and offset in x- and y-direction, see Table 8. The found approximation of the spatial accuracy of the computed offset revealed that the accuracy of repeated calibration was within 64m in x-direction and 26m in y-direction. Given the simple computation method of the overlay, this is quite accurate. However, 2.5 nest sites fit in the distance between the minimum and maximum in x-direction and one nest site in y-direction. For an

analysis differentiating between nest sites, this is inaccurate and unsatisfactory.

The final **snow coverage per nest site and image** was analysed using the presented RGB method instead of the also tested RGB normalized difference snow index (NDSI) [Hinkler et al. 2002]. The chosen RGB method is simple, its threshold value adaptable to the colour scheme and it performs accurately in neutral illumination conditions, such as in Figure 47b. The computation of the NDSI was impossible as the recording camera only recorded in RGB. Hinkler et al. 2002 presented two methods of an RGB NDSI to approximate the NDSI using only RGB photos. I decided to test their second method, which assumed homogeneously illuminated photos as the *2-Photos* suitable for analysis were not partially darkened. Furthermore, I overtook their empirically chosen snow threshold of 0.6. Indeed, the computed RGB NDSI overlooked snow less often than the RGB method on lightly bluish *2-Photos*, such as Figure 47a. On the other hand, this method falsely detected snow if there were trees near the camera covering parts of the detail. I assume that the trees influenced the analysis similarly as a partial cloud cover. Additionally, the RGB NDSI method falsely detected snow if the *2-Photo* was overexposed, such as Figure 47c. The here chosen camera took almost only overexposed *2-Photos* on sunny days, see above. To overcome most efficiently the shortages of either method, I chose to more rigorously select the suitable *2-Photos* for the RGB method, see below.

The main advantage of the function snow coverage per nest site and image was the automatic analysis of thousands of *2-Photos*. However, the current method requires between half a day and two days to analyse the footage of one camera of one year depending on the number of suitable *2-Photos*. In my original concept of the algorithm I planned to read in all *1-Models* at once and store them in R as the import of the *2-Photos* is the most time-consuming step. Only then, I would loop through the *2-Photos*. This would imply to import n + m images, where n is the number of *2-Photos* and m the number of *1-Models*. Unfortunately, R was unable to store more than five pictures at once in the used environment. Therefore, I imported the *1-Model* and the *2-Photo* every time they were needed. Consequently, I imported n * m images which led to immense program runtime. Eventually, no time was left to optimize a bug-free running algorithm. How I planned to improve the efficiency is explained in 4.5.

The selection of the *2-Photos* analysed for the **snow coverage per nest site and season** differed from the original selection of the suitable *2-Photos*. In the preparation, all *2-Photos* were classified as suitable or unsuitable for analysis based on the recorded weather condition. I tested the snow coverage computation on a sample and it produced reasonable results. However, the whole dataset contained *2-Photos* with extreme illumination conditions. The computed snow coverage per nest site and image of such *2-Photos* demonstrated that neither the RGB method nor the RGB NDSI method could cope under all conditions. Thus, I excluded *2-Photos* of extreme illumination conditions such as shadowy-blue or sunny-bright from the computation of the final snow coverage per season; The final snow coverage was computed for the neutral illuminated *2-Photos* only. Even after this additional selection process, most of the days were represented by at least one *2-Photo*. The lost data was approximately 1/8 of the suitable *2-Photos* at the Lago Cadagno in 2017. Yet, the range of the daily snow coverage decreased drastically. Still, some *2-Photos* of extreme illumination were overlooked. Especially the first snow-free day and last snow-covered day were gravely affected by such mistakes.

The statistical measure to smooth the resulting daily snow coverage, the 90% percentile, filtered the most severe cases of extreme illumination conditions, especially overexposure. However, it remained usually near the daily maximum. Furthermore, it differed often largely from the daily median. In general, the daily median showed a clear tendency for either the daily 10% percentile (rarely) or the daily 90% percentile (frequently). I conclude the 90% percentile is still affected by extreme values which probably originated in the extreme illumination conditions. Thus the presented thresholds and selections required a stronger smoothing than the 90% percentile. Yet, a too strong smoothing needs to be

avoided, as dirty or shadowy snow could be overlooked.

Visualising the seasonal snow coverage assisted largely the accuracy assessment and the comprehension of the computed data. It found a link between the weather and the snow coverage. I deduce, the computed data describes the general seasonal snow coverage. However, there was a massive difference between 10% and 90% percentile at the start of April in Figure 29 which could be explained as follows. Either, this variability of snow coverage is assumed to be true. This could be the case if a light snow fall occurred but melted away again within the subsequent day. Or, the difference between 10% and 90% percentile represent the noise of quantified snow coverage in all photos within the same day. This indicates the influence of the illumination. As there are hardly any days on the footage where the snow coverage melted within one day, let alone a repetition of such events, I excluded the first option. Thus, I expect the variation in snow coverage to correspond to the noise.

4.3 The statistical models

Three linear mixed effect models (LMEMs) analysed the **plausibility of the computed snow coverage variable**. While all models indicated a between-year variation of snow, its effects were strongest for first snow-free day and last snow-covered day. Possibly, the snow coverage of a year resembles one of at least three main types of snow coverage courses, see Figures 73 to 76. Then, the between-year variability would result from the onset and end off the snowmelt irrespective of how fast the snow melted. The slight correlation between the first snow-free day and the last snow-covered day, see Figure 65, support this view. Indeed, the current data indicates three main kinds of snow coverage courses for the Val Piora between 2015 and 2018, see 6.3.3. First, the constant snowmelt; Starting from a high snow coverage, followed by continuously melting, with few additional snowfall events, see Figure 75. Second, the repetition of the first kind of snowmelt cycles; with a various number of cycles in different years, see Figures 74 and 75. Third, the chaos. Here, the illustrated snow coverage course zigzags wildly. These severe fluctuations possibly originate in the wiggly cameras.

The three LMEMs investigated the effects of various methodological and environmental factors, see 2.5.1. Only environmental factors provided evidence of a strong effect in all models, i.e. strong effect means that the 95% credible interval of 2000 random simulations does not include 0 [F. Korner-Nievergelt and P. Korner-Nievergelt 2019]. This contributed to establishing proof that the three computed snow coverage variables first snow-free day, last snow-covered day and number of temporary snow-free days describe the snow coverage. On the other hand, methodological factors indicated evidence for an absence of a strong effect. The number of temporary snow-free days was affected by more methodological factors than the other two measures. Thus, it might describe the actual snow coverage less accurately.

The computed LMEM indicated evidence of strong effects of the median of slope, the median of elevation and the median of NDVI. Plains were modelled to be late snow-free for the first time whereas steep sites were early snow-free. Therefore, this result is in accordance with my expectation and the recorded snowmelt. Indeed, the *2-Photos* disclosed that snow does not remain on almost perpendicular sites. The last snow-covered day of the season was modelled earlier in season for low elevations which is in agreement with my expectation. The footage provided additional evidence for this finding; The last snowfall of the summer season did not reach the bottom of the valley anymore and only added to the snow coverage at positions of higher altitudes. Steep slopes were unsuited for a snow to accumulate and hence fresh snow melted away within days. Additionally, the site was numerous days temporary snow-free if the median normalized difference vegetation index (NDVI) was high. Possibly, snow remained undetected between dense vegetation or it melted quicker than on sparsely vegetated ground because of an influence of the land cover on the microclimate. Despite these mostly comprehensible and expected results, they are affected by methodological factors and random noise. The random noise may origin from the temporal, spatial and spectral inaccuracies of the snow coverage algorithm.

Two generalized linear model (GLM) investigated the **influence of environmental factors on the nest site choice**. To create a robust model fitting my data, I created the GLM without snow, equation 14. Most of the modelled preferences were in accordance with the existing literature. Thus, the model might add to the understanding of the nest site choice of the northern wheatear in the Alps. Despite a few differences between the modelled preference and the data, the Val Piora appears as a suitable habitat in theory and in reality.

The GLM predicted an altitude preference of 1900m a.s.l. to 2100m a.s.l. My expectation of a preference of lower altitudes is in agreement with this prediction. On the other hand, the reported population

maximum occurs in more elevated areas, 2200m a.s.l. to 2500m a.s.l. [Knaus et al. 2018]. The flat parts of the study area remained mostly below 2100m a.s.l. which might explain the found mismatch between the model and the distribution of the Swiss population. Furthermore, Glutz von Blotzheim 1962 described a preference for nest sites at and above the treeline. The eastern part of the Val Piora only accommodates a few trees and consequently appears to be located above the treeline. However, most nest sites in the Val Piora are located below the typical altitude for the tree line in the southern Alps at ~2200m a.s.l. [Paulsen and Körner 2001]. The natural treeline is commonly described as a specific elevation and often influenced by alpine farming. Possibly, the absence of trees influences the nest site choice stronger than the elevation of the natural treeline.

The nest sites were predicted as slopes between 15° and 30° with little variation in their inclination. This is in line with my expectation. The modelled preference and the maximum of the median of slope agrees with Wartmann 1985. The maximum of the variable median of slope of a nest site at Val Piora was 49.6°.

The estimates of the model suggested that wheatears display a preference for nest sites with aspects S to SW. Such a preference for the sunny slopes is in accordance with my expectation and Wartmann 1985.

The model predicted a preference for even to slightly concave grounds. However, I expected roughness values which indicated cairns or other narrow and convex features. Either, this mismatch between the model and my expectation might origin in the roughness raster. The raster indicated no cairns at any known location of these piles of stones. Possibly, the kernel radius used to compute the roughness from the DEM was not ideal. Additionally, the cairns might not have been detected because of the coarse spatial resolution of the DEM. Without the knowledge of their position, I would have been unable to detect them on the DEM myself. Consequently, the used data basis prevented a fine-scaled description of the existing roughness. Or, the cairns and other convex structures are less preferred for nesting than for hunting. Furthermore, concave hillsides might shield the nest better from predators and weather than convex structures. A combination of both explanations might also be possible.

The estimates for the median of the NDVI indicated a preference for bare ground or rocks and sparse vegetation. The predicted land covers are in accordance with my expectation and the literature [Conder 1989; Hussell et al. 2014; Tye 1992; van Oosten et al. 2014]. However, a careful direct comparison of the results and the literature is in order as my result is NDVI-based while the literature has been based on manually sampled land cover.

The land covers within a nest site were predicted to range. These heterogeneous land covers, indicated by a large 90% percentile distance in NDVI, are in accordance with my expectation. On the other hand, the data displayed less heterogeneity. Such uniform NDVI values were mostly found on the flat valley bottom. This could indicate that flat slopes might be preferred over sparsely vegetated areas, as long as suitable hunting grounds are near.

The model describes a preference of proximity to the three nearest nests. No literature values of CD3N exist to my knowledge. However, small CD3N indicate a preference of nesting in small clusters. Such small clusters are in accordance with Arlt and Part 2007 who recorded 70% of all found territories in "small clusters of 2-5 pairs". On the other hand, the model found no limit for too near nests, which I expected. To find such a limit, the CD3N should have been included as a quadratic term into the GLM. However, this introduced model fitting difficulties which prohibited the inclusion of this term.

Similar to the CD3N, small values were preferred and the expected limit for small distances lacked. Again, the required inclusion of the quadratic term of the distance to the nearest path prevented the model from converging.

The GLM with snow, see equation 15, considered the effects of some of the most influential environmental factors of the GLM without snow (median of the aspect, median of the elevation, median of the NDVI, median of the roughness, median of the slope, cumulated distance to the three nearest nests and interaction of the median of elevation with the median of slope) and the effects of the snow coverage (first snow-free day and last snow-covered day) on the nest site selection. As is the model without snow, slope displayed a strong effect suggesting that wheatears preferred slopes with inclinations between 25° and 30°. Furthermore, the model hinted at a preference for nest sites whose first snow-free day is early, e.g. before April. This is in accordance with my expectation. Yet, the result of the snow coverage algorithm suggested later dates, e.g. mid- to end of May, which is shortly after the arrival of the northern wheatears. Moreover, the model indicated a preference for nest sites with a last snow-covered days later in season, e.g. end of May. On the other hand, the data indicates slightly earlier dates of mid-May to end of May. A preference of snow-free sites by the end of May was partly in accordance with Wartmann 1985. To be precise, he compared the supply and utilisation of areas [%] [Wartmann 1985]. He found an utilisation 3.2% higher than supply for areas snow-free before the 14th of May and a 6.3% higher utilisation for sites snow-free between the 14th of May and the 31th of May [Wartmann 1985]. Furthermore, the utilisation of areas snow-free between the 1st of June and the 22nd of June was 6.6% below the supply and sites snow-free after the 22nd of June were 2.9% less used than supplied. In the Val Piora, only a few sites are still snow-covered by the end of May. Thus, a thorough comparison of the model to the reality or Wartmann 1985 is not possible. Additionally, the effects of the snow coverage variables were uncertain and small compared to the effect of the median of slope. The residuals displayed not only a more abrupt detachment of the diagonal in the Q-Q-plot but also a larger leverage than the residuals of the GLM without snow, see Figure 85.

The presented GLM with snow was created without the variables number of temporary snow-free days, observation year, distance to nearest path and others. The above results are less reliable than the results of the GLM without snow because of the encountered converge difficulties, the unexplained spatial autocorrelation and the unscrutinised between-year variation. Both GLMs did not analyse the between-year variation. Other than in the GLM without snow, the exclusion of the observation year cannot be justified in this model. On the contrary, each of the above LMEMs indicated a strong between-year variation for the first snow-free day and the last snow-covered day. Consequently, the GLM with snow could vary between years. Additional caution in the interpretation demands the exclusion of the distance to the nearest path. The spatial autocorrelation found in the GLM without snow was explained by the distance to the nearest path. Thus, the GLM with snow might be spatially autocorrelated. Number of temporary snow-free days could not be fitted despite various data transformations. Possibly, this variable was zero-inflated [F. Korner-Nievergelt, Roth, et al. 2015], see Figure 65.

4.4 The conclusions

4.4.1 Contributions

The fitted GLM without snow appears to predict nest sites on S- to SW-exposed slopes; elevations between 1900m a.s.l. and 2100m a.s.l.; bare ground and sparse vegetation; ranges in land cover; even to slightly concave grounds; 15° to 30° steep slopes; little variation in the incline of these slopes; sites close to the three nearest nests and sites close to the nearest path. The analysed environmental factors appear to indicate no between-year variation in the present data. Additionally, the GLM with snow seems to predict a preference for early first snow-free days, i.e. before April and late last snow-covered days, i.e. end of May.

The distribution of dates of the first snow-free day and the last snow-covered day indicate a large between-year variance of snow coverage. The first snow-free day seems to be later for flatter slopes. The last snow-covered day appears to be earlier in lower altitudes. More temporary snow-free days are indicated on steeper slopes and more densely vegetated areas. Most of Val Piora is snow-free before the end of May. The seasonal snow coverage might possibly be generalised either as constant decrease, as repeated cycle of constant decreases or as chaos.

The minimal distance between two nests was 15m in 2013. The chosen radius of a nest site (25m) approximates the average distance between the two nearest nests per year. The data indicates nests in small clusters. The rocky beaches of the Lago Ritom are used as nest sites by the northern wheatear.

The chosen threshold RGB(0.6, 0.6, 0.75) appears to accurately detect snow in neutral illuminations and could be optimized for other colour schemes. The recorded *Photos* depict almost the entire summer season. The three main illumination conditions of the suitable *2-Photos* are bluish-shadowy, neutral and bright-sunny. The overlay function can focus on different areas of interest. A low spatial accuracy is achieved with the simple overlay function. Mounting the cameras to trees could have contributed to the spatial variation in the recorded detail (wiggling camera). The RGB NDSI method tolerates lightly bluish illumination conditions. The automation of snow coverage per nest site and image facilitates the analysis of thousands of *2-Photos*. The seasonal snow coverage is fast computed and computes snow coverage as the 90% percentile per day. The slight smoothing of the 90% percentile filters the most extreme values. The chosen statistical measure influences the result of snow coverage per nest site and season largely. The computed snow coverage regularly seems to increase after a snowfall which was clearly visible in the climatogram. The snow coverage variables might contribute to the description of the snow coverage rather than the used method.

The creation of remotely sensed variables for the exploration of the effects of environmental factors is relatively fast and could be effortlessly applied on larger study areas. The NDVI recorded at the end of August might approximate the ratio of vegetated or unvegetated grounds of the early summer.

4.4.2 Limitations

The nests on the rocky beaches of the dammed up Lago Ritom appear to be in or under the water on the DEM. The study area remains mostly below 2100m a.s.l., below the elevation of the maximal Swiss population density. The spatial resolution of the DEM appears too coarse to compute a roughness fine-scaled enough to detect cairns. The late sampling date of the NDVI shows not the vegetation encountered on the arrival of the northern wheatears in spring. The NDVI raster is computed for the year 2015 only. The computed NDVI seems less detailed than the recorded land cover data. The prioritising of the environmental factors affecting the nest site choice remains unclear. The visual assessment of clustering of nests was not statistically tested. The influence of the various density measures and spatial scales was not analysed. The preference of path proximity might hint at a sampling bias in geographic space. The absence points are pseudo-randomly created and lack the sampled land cover data. During the overlay, complete congruence of the *2-Photo* and the *1-Model* is not achieved. The recorded scenes vary slightly over time preventing the analysis of the same geographic area. Not all nests were recorded or their extent was not depicted completely. The timestamp of the *Photos* becomes less accurate towards the end of the footage. The used cameras lack a near infrared channel. *2-Photos* taken on sunny days tend to be overexposed. Certain *2-Photos* display a dark halo around the centre. Extreme illumination conditions exist for a multitude of *2-Photos*. A single RGB threshold is insufficient to detect snow under all illumination conditions and to reliably distinguish snow from other surfaces. The RGB NDSI method falsely detected snow if trees covered parts of the scene or if the *2-Photos* are overexposed. The choice of GCPs is restricted to features depicted on the *2-Photo* and the *1-Model*. The 90% percentile seems to smooth extreme values too little. The first snow-free day and last snow-covered day are severely affected by a single day of faulty computed snow coverage. The developed snow coverage algorithm performs inefficiently and requires much manual preparation. The first snow-free day, last snow-covered day and number of temporary snow-free days are influenced by methodological factors and random noise.

The GLMM and GLM with snow encounters immense fitting difficulties. Variables, such as the observation year and the distance to the nearest path, were excluded from the GLM with snow because of these problems. The GLM without snow still faces some difficulties with the distance measures; Neither CD3N nor the distance to the nearest path could be included as quadratic terms. The data set to fit the GLM with snow was small (69 nest sites). The GLM with snow indicated an uncertain and small effect of the snow coverage. The possibility of spatial autocorrelation can not be excluded from the GLM with snow. The between-year variability could not be scrutinised thoroughly because of the static remotely sensed data and the small dataset reporting the snow coverage.

4.5 An outlook

The used remotely sensed raster data contributed largely to the findings of the statistical models. For further investigations, I suggest the use of a DEM with a high spatial resolution (<0.5m cell size), especially for the computation of the roughness. This resolution should allow the detection of cairns, maybe an optimization of the used kernel radius is required. To scrutinise the effect of higher elevation, I suggest to search for northern wheatears nests in higher altitudes around the Val Piora or in the Alps and to include these areas into the study area. To include the nests on the shores of the Lago Ritom, I recommend a DEM which was recorded when the dam was almost empty.

The NDVI described the land cover of the study area and varies over the course of the season. The nest site selection process of the northern wheatear partly happens after the breeding season and partly at the time of territory establishment [Arlt and Pärt 2008a]. For the nest site selection process during the territory establishment, a change of the recording date is advisable for the false coloured infrared orthophotomosaic (FCIR) from which the NDVI was computed. This FCIR should be recorded during the breeding period of the northern wheatear; The optimal recording date was during the time of nest-building, towards mid-May. To consider the between-year variability of the vegetation during the territory establishment respectively after the breeding season, I recommend to calculate these NDVI rasters using FCIR rasters recorded in mid-May and in August in the particular year. I expect such yearly NDVIs to approximate the sampled land cover more detailed. I further advise to sample the land cover also for absence points, i.e. locations where never before a nest was. Almost all nests close to path are assumed to be detected. To reach the same certainty for nests further away from the paths, I suggest to also patrol parallel to the paths. This additional data permits to investigate the preference of path proximity more thoroughly. Furthermore, I propose to statistically show the clustering of the nests, e.g. by comparing the found distribution to a random distribution [Clark and Evans 1954].

Further not yet satisfyingly answered questions concern additional environmental factors and their prioritising in the nest site choice of the northern wheatear: Is the elevation (as an approximation for the natural treeline) or the mere absence of trees more important? Does the distance to water bodies or wet grounds affect the nest site choice? Is the steepness of the ground or the heterogeneous land cover (approximated here by the 90% percentile distance of the NDVI) more influential on the nest site choice? What are the environmental factors affecting the hunting grounds of the northern wheatears in Val Piora? Where are these hunting grounds? Does the distance to the hunting grounds affect the nest site choice of the northern wheatears?

The presented recorded *Photos* and the developed snow coverage algorithm requires improvement. The following adaptions could add to a more efficient and more accurate description of the snow coverage course in the Val Piora. I suggest a different, not R-bound, georeferencing method to increase the spatial accuracy of the of the overlay function and the location of the nest sites on the *2-Photos*. Furthermore, I advise to install fixed GCPs in Val Piora, such as black-and-white crosses which are visible for the camera. The cameras ought to be stably mounted for this automated analysis method. If possible, they should be attached to poles or the lower part of a thick tree trunk. The precise location of the camera in x/y/z could add to the creation of congruent *1-Models*. The documentation of snow coverage in this alpine valley requires the use of cameras which are cold-resistant, record the actual time, correct for overexposure, correct for the blue hour and create evenly exposed *2-Photos*. To improve the detection of snow at the nest sites, I suggest to test the second RGB NDSI computation method by Hinkler et al. 2002, which takes the overall level of brightness into account instead of the mean RGB. This method might more robustly compute the snow coverage under all illumination conditions or it might require the exclusion of the *2-Photos* with the most extreme illumination conditions only. The rest of the workflow

could remain unaltered. On the other hand, if the RGB method should be used again on the existing *2-Photos*, I propose the use of three different RGB thresholds of snow - one for blusih-shadowy, one for neutral illuminated and one for sunny-bright *2-Photos*. This also requires the further division of the suitable *2-Photos* into these three categories and the according adaptions of the algorithm. After these proposed changes, the 90% percentile might describe the daily snow coverage correctly. Nonetheless, it was worth investigating further if the daily median or the daily 90% percentile was the more suitable statistical measure.

To document all nest sites completely, I recommend the setup of more cameras or an airborne recording method. The resulting large dataset might add to a better fit of a GLMM. Ideally, a FCIR raster of the entire Val Piora is recorded each day allowing the computation of the NDSI to compute the snow coverage.

To improve the efficiency of the algorithm, I propose an additional analysis step before the snow coverage per nest site and image commences. I assume that all *1-Models* of one camera and year, which I exported from ArcScene, are congruent pixel-wise and thus interchangeable; Their only difference is the location of the red area. The additional method imports one *1-Model* at a time, crops it, creates the red-pixel-index and stores this index together with the nest ID in *9-Index* (to be created). Then, the snow coverage per nest site and image could import any one of the *1-Models*. While the program loops through all *2-Photos*, it uses the stored *1-Model* for the resampling. Instead of computing the red-pixel-index each time, it could be looked up in *9-Index*. Otherwise, the snow coverage could have been analysed as described above. I predict, this method reduces the calculation time dramatically.

The computed statistical models provided evidence towards an understanding of the effect of the environmental factors on the nest site choice of the northern wheatears in the Val Piora. However, the limited snow coverage data and the static vegetation data prohibited to assess the influence of the between-year variation of the environmental data. The above proposed changes in the NDVI, land cover sampling and snow coverage computation could allow a thorough investigation. I recommend to optimize the computed snow coverage until it produces a robust and certain measure for all nest sites. Possibly, these alterations also reduce the methodological effects detected by the LMEM on the computed snow coverage variables.

Further optimizations of the models might include other data transformation methods, higher polynomial terms or an informative prior. I propose to log-transform the distance measures and to model at least their linear and quadratic terms. Instead of a GLMM, a GLM without snow and a GLM with snow, I suggest to create one GLMM only, which contains the computed robust snow coverage for each nest site, the land cover values, the distance to the nearest path and the observation year. Thus, I expect a more thorough investigation of the between-year variation of the environmental factors snow and vegetation while also recreating the possibly still existing spatial autocorrelation. The such created GLMM could then be used to predict a suitability map for nest sites in the Val Piora, such as Figure 71.

5 REFERENCES

Arizaga, J., & Bairlein, F. (2011). The role of age, sex, subspecies, body size and fuel load in determining the social status of a migratory passerine during the non-breeding period. *Ornis Fennica, 88*(3), 154–160.

Arlt, D., & Part, T. (2007). Nonideal breeding habitat selection: a mismatch between prefreence and fitness. *Ecology, 88*(3), 792–801.

Arlt, D., & Pärt, T. (2008a). Post-breeding information gathering and breeding territory shifts in northern wheatears. *Journal of Animal Ecology, 77*, 211–219. doi:10.1111/j.1365-2656.2007.01329.x

Arlt, D., & Pärt, T. (2008b). Post-breeding information gathering and breeding territory shifts in northern wheatears - Appendix. *Journal of Animal Ecology, 77*, 17.

Bachofen, R., Barudoni, F., Dolfini, A., Guscetti, G., Jelmini, V., Loizeau, J.-l., ... Wildi, W. (2016). *Piora - Lago di Cadagno - Ritom. Natur- und Umwelfuehrer*. Fondazione Centro Biologia Alpina. Bellinzona.

Bavay, M., Grünewald, T., & Lehning, M. (2013). Response of snow cover and runoff to climate change in high Alpine catchments of Eastern Switzerland. *Advances in Water Resources, 55*, 4–16. doi:10.1016/j.advwatres.2012.12.009

Birdlife International. (2018). Northern Wheatear - Oenanthe oenanthe. BirdLife International.

BirdLife Schweiz. (2018). Tweet from 15.11.2018.

Bormann, K. J., Brown, R. D., Derksen, C., & Painter, T. H. (2018). Estimating snow-cover trends from space Snow monitoring from space. *8*(November). doi:10.1038/s41558-018-0318-3

Bötsch, Y., Tablado, Z., & Jenni, L. (2017). Experimental evidence of human recreational disturbance effects on bird-territory establishment. *Proceedings of the Royal Society B: Biological Sciences, 284*(1858). doi:10.1098/rspb.2017.0846

Bötsch, Y., Tablado, Z., Scherl, D., Kéry, M., Graf, R. F., & Jenni, L. (2018). Effect of Recreational Trails on Forest Birds: Human Presence Matters. *Frontiers in Ecology and Evolution, 6*(November), 1–10. doi:10.3389/fevo.2018.00175

Brooke, M. d. L. (1979). Differences in the Quality of Territories Held by Wheatears (Oenanthe oenanthe). *The Journal of Animal Ecology, 48*(1), 21–32. doi:10.2307/4097

Bundesamt für Umwelt BAFU. (2018). Murinascia Grande. Retrieved June 19, 2019, from https://api3.geo.admin.ch/rest/services/ech/MapServer/ch.bafu.wasser-teileinzugsgebiete%7B%5C_%7D2/51116/extendedHtmlPopup?lang=de

Bürki, P. (2014). *Einfluss des Geländes auf ökologische Eignungsanalysen* (Doctoral dissertation).

Catzeflis, F. (1976). Les oiseaux nicheurs du col de balme (Trient, VS). *Bulletin de la Murithienne, 92*, 81–92.

Chamberlain, D. E., Negro, M., Caprio, E., & Rolando, A. (2013). Assessing the sensitivity of alpine birds to potential future changes in habitat and climate to inform management strategies. *Biological Conservation, 167*, 127–135. doi:10.1016/j.biocon.2013.07.036

Chefaoui, R. M., & Lobo, J. M. (2008). Assessing the effects of pseudo-absences on predictive distribution model performance. *Ecological Modelling, 210*(4), 478–486. doi:10.1016/j.ecolmodel.2007.08.010

Clark, P. J., & Evans, F. C. (1954). Distance to Nearest Neighbor as a Measure of Spatial Relationship in Populations. *Ecology, 35*(4), 445–453.

Conder, P. (1989). *The Wheatear*. London: Christopher Helm.

Cooley, S. (2016). Terrain Roughness - 13 Ways. Retrieved from http://gis4geomorphology.com/roughness-topographic-position/

Corti, U. A., & Melcher, R. (1958). Beiträge zur Kenntnis der Vogelwelt Graubündens, 2. Periode. *Jahres-bericht der Naturforschenden Gesellschaft Graubünden, 87 (1957-1,* 72–106.

Cramp, S., Brooks, D. J., Hall-Craggs, J., Ogilvie, M. A., Simmons, K. E. L., Wilson, M. G., ... Wallace, D. I. M. (1988). *Handbook of the birds of Europe, the Middle East and North Africa. Volume V - Tyrant flycatchers to thrushes.* Oxford Univ. Press.

Currie, D., Thompson, D. B. A., & Burke, T. (2000). Patterns of territory settlement and consequences for breeding success in the Northern Wheatear Oenanthe oenanthe. *Ibis, 142,* 389–398. doi:10.1111/j.1474-919X.2000.tb04435.x

Duguay, C. R., & LeDrew, E. F. (2006). Mapping Surface Albedo in the East Slope of the Colorado Front Range, U.S.A., with Landsat Thematic Mapper. *Arctic and Alpine Research, 23*(2), 213. doi:10.2307/1551385

Finney, S. K., Pearce-Higgins, J. W., & Yalden, D. W. (2005). The effect of recreational disturbance on an upland breeding bird, the golden plover Pluvialis apricaria. *Biological Conservation, 121*(1), 53–63. doi:10.1016/j.biocon.2004.04.009

Fretwell, S. D., & Lucas, H. L. J. (1969). On territorial behaviour and other factors influencing habitat distribution in birds. *Acta Biotheoretica, 19*(1), 16–36. doi:10.1007/BF01601953

Glutz von Blotzheim, U. N. (1962). *Die Brutvögel der Schweiz: eine Zusammenfassung unserer heutigen Kenntnisse über Verbreitung, Bestandesdichte, Ernährung und Fortpflanzung der seit 1900 in der Schweiz als Brutvögel nachgewiesenen Arten.* Aarau: Verlag Aargauer Tagblatt AG.

Glutz von Blotzheim, U. N., & Bauer, K. M. (1988). *Handbuch der Vögel Mitteleuropas. Band 11/I - Passeriformes (2. Teil).* Wiesbaden: AULA-Verlag.

Graham, C. H., Ferrier, S., Huettman, F., Moritz, C., & Peterson, A. T. (2004). New developments in museum-based informatics and applications in biodiversity analysis. *Trends in Ecology and Evolution, 19*(9), 497–503. doi:10.1016/j.tree.2004.07.006

Hågvar, S., Glesne, O., & Østbye, E. (2009). Food habits and niche overlap in three alpine passerine birds, South Norway. *Ornis Norvegica, 32*(1), 56–73.

Heath, M. F., Borggreve, C., Peet, N., & Hagemeijer, W. J. (2000). *European bird populations: Estimates and trends.* Cambridge: BirdLife International.

Hinkler, J., Pedersen, S. B., Rasch, M., & Hansen, B. U. (2002). Automatic snow cover monitoring at high temporal and spatial resolution, using images taken by a standard digital camera. *International Journal of Remote Sensing, 23*(21), 4669–4682. doi:10.1080/01431160110113881

Hussell, D. J., Bairlein, F., & Dunn, E. H. (2014). Double brooding by the Northern Wheatear on Baffin Island. *Arctic, 67*(2), 167–172. doi:10.14430/arctic4387

Kanagaraj, R., Wiegand, T., Mohamed, A., & Kramer-Schadt, S. (2013). Modelling species distributions to map the road towards carnivore conservation in the tropics. *Raffles Bulletin of Zoology,* (SUPPL.28), 85–107. doi:10.1111/j.2041-210X.2011.00172.x. arXiv: 1132

Knaus, P., Antoniazza, S., Wechsler, S., Guélat, J., Kéry, M., Strebel, N., & Sattler, T. (2018). *Schweizer Brutvogelatlas, 2013 - 2016. Verbreitung und Bestandsentwicklung der Vögel in der Schweiz und im Fürstentum Liechtenstein.* Sempach: Schweizerische Vogelwarte Sempach.

Korner-Nievergelt, F., & Korner-Nievergelt, P. (2019). Personal Conversation.

Korner-Nievergelt, F., Roth, T., von Felten, S., Guélat, J., Almasi, B., & Konrer-Nievergelt, P. (2015). *Bayesian Data Analysis in Ecology Using Linear Models with R, BUGS, and Stan.*

Lillesand, T. M., Kiefer, R. W., & Chipman, J. W. (2008). *Remote Sensing and Image Interpretation* (6th ed.).

Maggini, I., & Bairlein, F. (2012). Innate sex differences in the timing of spring migration in a songbird. *PLoS ONE, 7*(2), 5–8. doi:10.1371/journal.pone.0031271

Mattes, H., Maurizio, R., & Bürkli, W. (2005). *Die Vogelwelt im Oberengadin, Bergell und Puschlav ein Natur-führer zur Avifauna in einem inneralpin*. Sempach: Schweizerische Vogelwarte Sempach.

Maumary, L., Vallotton, L., & Knaus, P. (2007). Steinschmätzer. In *Die vögel der schweiz* (Chap. Steinschmä, pp. 568–571). Schweizerische Vogelwarte, Sempach und Nos Oiseaux, Montmollin.

Meier, C. (2018). Personal Conversation.

Meier, C., Weibel, R., & Liechti, F. (2018). Personal Conversation.

Messmer, N. (2019). Personal Conversation.

MeteoSchweiz. (2016). Der Hitzesommer 2015 in der Schweiz. (260), 68.

Öberg, M. (2014). *When Climate is Changing: Effects of Phenology and Local Climate on Individual Fitness*.

Olliver, P., Debout, C., & Debout, G. (1999). Importance du choix du territoire dans la reproduction du traquet motteux Oenanthe oenanthe sur une dune fixée de la manche (N-O France). *Alauda*, *67*(3), 213–222.

Pärt, T. (2001). The effects of territory quality on age-dependent reproductive performance in the northern wheatear, Oenanthe oenanthe. *Animal Behaviour*, *62*(2), 379–388. doi:10.1006/anbe.2001.1754

Pärt, T., Knape, J., Low, M., Öberg, M., & Arlt, D. (2017). Disentangling the effects of date, individual, and territory quality on the seasonal decline in fitness. *Ecology*, *98*(8), 2102–2110. doi:10.1002/ecy.1891

Paulsen, J., & Körner, C. (2001). GIS analysis of treeline elevation in the Swiss Alps suggests no exposure effect. *Journal of Vegetation Science*, *12*(6), 817–824. doi:10.2307/3236869

Pollheimer, J., & Gstir, J. (2013). Abundance and habitat selection of two alpine songbird species in the National Park Hohe Tauern (Austria). *5*(June), 601–602.

QGIS. (n.d.). A gentle introduction in GIS. Retrieved June 21, 2019, from https://docs.qgis.org/2.8/en/docs/gentle%7B5C_%7Dgis%7B5C_%7Dintroduction/vector%7B5C_%7Ddata.html

Sánchez-Bayo, F., & Wyckhuys, K. A. (2019). Worldwide decline of the entomofauna: A review of its drivers. *Biological Conservation*, *232*(September 2018), 8–27. doi:10.1016/j.biocon.2019.01.020

Scherrer, S. C., Appenzeller, C., & Laternser, M. (2004). Trends in Swiss Alpine snow days: The role of local- and large-scale climate variability. *Geophysical Research Letters*, *31*(13). doi:10.1029/2004GL020255

Schmid, H., Luder, R., Naef-Daenzer, B., Graf, R., & Zbinden, N. (1998). *Schweizer Brutvogelatlas Verbreitung der Brutvögel in der Schweiz und im Fürstentum Liechtenstein*. Sempach: Schweizerische Vogelwarte Sempach.

Schrauth, F. E., & Wink, M. (2018). Changes in species composition of birds and declining number of breeding territories over 40 years in a nature conservation area in southwest Germany. *Diversity*, *10*(3). doi:10.3390/d10030097

Schweizerische Vogelwarte Sempach. (n.d.). Steinschmätzer. Retrieved April 19, 2019, from https://www.vogelwarte.ch/de/voegel/voegel-der-schweiz/steinschmaetzer

Sumser, H., Stenmans, W., Jongejans, E., Schwan, H., Siepel, H., Goulson, D., ... Hallmann, C. A. (2017). More than 75 percent decline over 27 years in total flying insect biomass in protected areas. *Plos One*, *12*(10), 1–13. doi:10.1371/journal.pone.0185809

Svensson, L., Mullarney, K., & Zetterström, D. (2009). Steinschmätzer. In *Der kosmos vogelführer* (282f.).

Swisstopo. (2019). Teileinzugsgebiete 2km2. Retrieved May 3, 2019, from https://api3.geo.admin.ch/rest/services/ech/MapServer/ch.bafu.wasser-teileinzugsgebiete%7B5C_%7D2/1956/extendedHtmlPopup?lang=de

Swisstopo. (n.d.[a]). swissALTI3D. Retrieved May 17, 2019, from https://shop.swisstopo.admin.ch/de/products/height%7B5C_%7Dmodels/alti3D

Swisstopo. (n.d.[b]). SWISSIMAGE FCIR. Retrieved May 17, 2019, from https://shop.swisstopo.admin.ch/de/products/images/ortho%7B%5C_%7Dimages/SWISSIMAGE%7B%5C_%7DFCIR%7B%5C#%7Dfurther-info

Swisstopo. (n.d.[c]). swissTLM3D. Retrieved May 17, 2019, from https://shop.swisstopo.admin.ch/de/products/landscape/tlm3D

Tye, A. (1992). Assessment of territory quality and its effects on breeding success in a migrant passerine, the Wheatear Oenanthe oenanthe. *IBIS*, *134*, 273–285.

van Oosten, H. H., van den Burg, A., Versluijs, R., & Siepel, H. (2014). Habitat selection of brood-rearing Northern Wheatears Oenanthe oenanthe and their invertebrate prey. *Ardea*, *102*(1), 61–69. doi:10.5253/078.102.0111

Wartmann, B. A. (1985). *Vergleichende Untersuchungen zur Populations-, und Nahrungsökologie von Wasserpieper und Steinschmätzer im Dischmatal GR* (Doctoral dissertation, University of Zurich).

Winkler, R. (1999). *Avifauna der Schweiz*. Der Ornithologische Beobachter.

6 APPENDIX

6.1 The used software

- Adobe Photoshop CC 2018, 8.1 6.3.9600.17415
- ArcMap, 10.5.1
- ArcScene 10.5.1
- Dropbox
- GeoGebra
- Google Drive
- Microsoft Access, 1901
- Microsoft Excel, 1901
- Microsoft Word, 1901
- MiKTeX, 2.9
- Notepad++, 7.5.9
- Latex stylesheet, uzhpub.cls, Copyright 2010 Juergen Spitzmueller
- Lucidchart
- R, version 3.4.4, used packages:
    - arm, 1.10.1
    - blmeco, 1.2
    - dplyr, 0.7.5
    - gstat, 1.1.6
    - ggplot2, 3.1.0
    - ggspatial, 0.2.1
    - grDevices, 3.4.4
    - gridExtra, 2.3
    - jpeg, 0.1.8
    - lubridate, 1.7.4
    - magick, 2.0
    - purrr, 0.2.4
    - readr, 1.1.1
    - raster, 2.6.7
    - rasterVis, 0.45
    - sp, 1.2.7
    - stringr, 1.3.0
    - tibble, 1.4.2
    - tidyr, 0.8.0
    - zoo, 1.8.1
    - tibble, 1.4.2
- TeXstudio, 2.12.14
- QGIS Desktop, 3.6.1 with GRASS 7.6.1

6.2 The additional data to methods

6.2.1 Accuracy of nest site location

I decided to quantify the difference between the true position of a nest and the recorded coordinates. Hence, I counted the steps between the entrance of the nest as depicted in Figure 48c and the GPS signal on Friday, 2018/08/03, see Table 9.

(a) The nest site.   (b) The immediate surroundings of the nest.   (c) The location to the nest entry.

Figure 48: The photos were taken at the Lago Ritom on the 13/06/2017.

The Swiss Ornithological Institute (SOI) documented in photos the nest site, the immediate surroundings and the nest entry, see 48. This documentation allowed to compare the true nest location to the recorded GPS position, see 6.2.1. On average, the GPS positioned the nest 3m away from the photographic documented location. The minimal difference between the two locations was 1m and the maximal difference 11m. These differences in position are relatively small given the surrounding mountains and the used GPS device.

The nest locations were recorded in QField on Android tablets using the inbuilt GPS device. Such navigation devices are thought to be imprecise, especially in areas with trees, buildings or mountains blocking the satellite signals. Last is the case in Val Piora. These spatial errors affected only the variables I recorded concerning cavities and perches as I was guided solely by the GPS location to the nests. These variables might not always document the correct location but they overlap with the nest site: the radius of the nest site was 25m and the maximal displacement only 11m. The work of the field assistants remained unbiased as they memorized the true nest location.

A more influential spatial error is the difference between the recorded coordinates and the true coordinates of this point. The accuracy of a GPS device in a tablet is often a few metres. The precise accuracy depends not only on the device but also on the availability of satellites. This spatial error is assumedly larger than the spatial errors of the raster layers DEM, slope, aspect, roughness and NDVI. However, I did not assess this accuracy in the Val Piora.

NestID	Year	Difference [Step]
A55	2018	8
A30	2018	1
C52	2018	2
A05	2018	6
A36	2018	4
A22	2018	5
B43	2018	15
C38	2018	5
B27	2018	2
A54	2018	4
A45	2018	3
A43	2018	2
C50	2018	2
B12	2018	4
B26	2018	3
B13	2018	7

Table 9: The minimal difference between the GPS destination and the true nest location was one step, which is approximately 70cm. The maximal difference was 15 steps (10.5m). On average 4.5 steps (3.1m) laid between the two points.

6.2.2 Land cover

The fields assistants estimated the land cover of a nest site at different points in time. They recorded the proportions of the six land cover classes (1) vegetation <5cm, (2) vegetation 5 - 15cm, (3) vegetation >15cm, (4) shrubs, (5) bare ground and (6) rocks per nest site, see Figure 49.
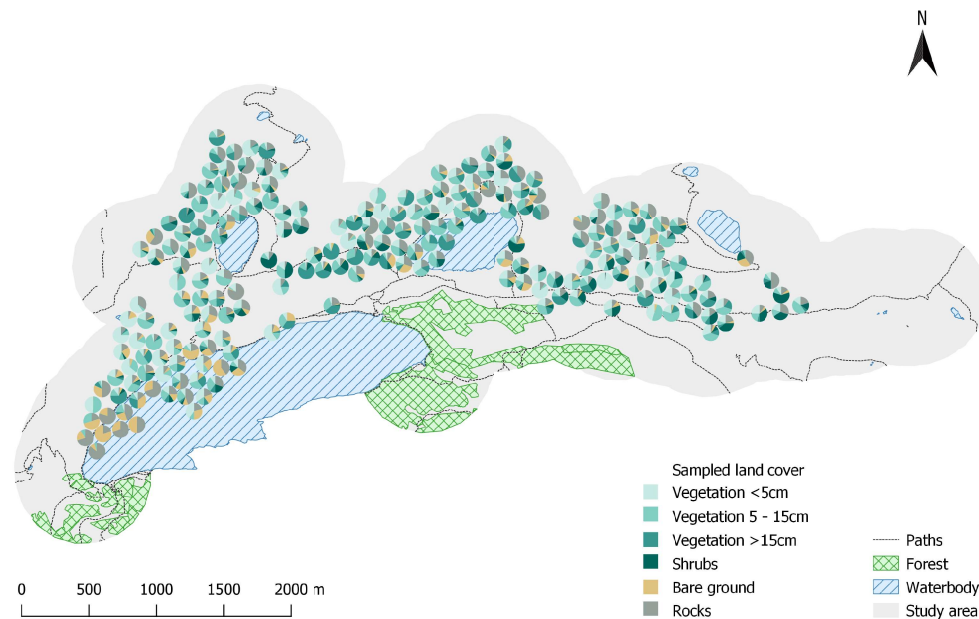

Figure 49: The pie charts reveal the sampled land cover compositions per nest site. Nest sites at the Lago Ritom, especially in the western part, disclose distinctively high ratios of bare ground and rocks. Nest sites with large proportions of shrubs are located along the Murinascia Grande and between the Lago Cadagno and Lago di Tom.

The land cover, particularly the vegetation, is dynamic and varies not only between years but also during the course of the year, see Figure 50. The sampling dates ranged between 15th May and 28th July.
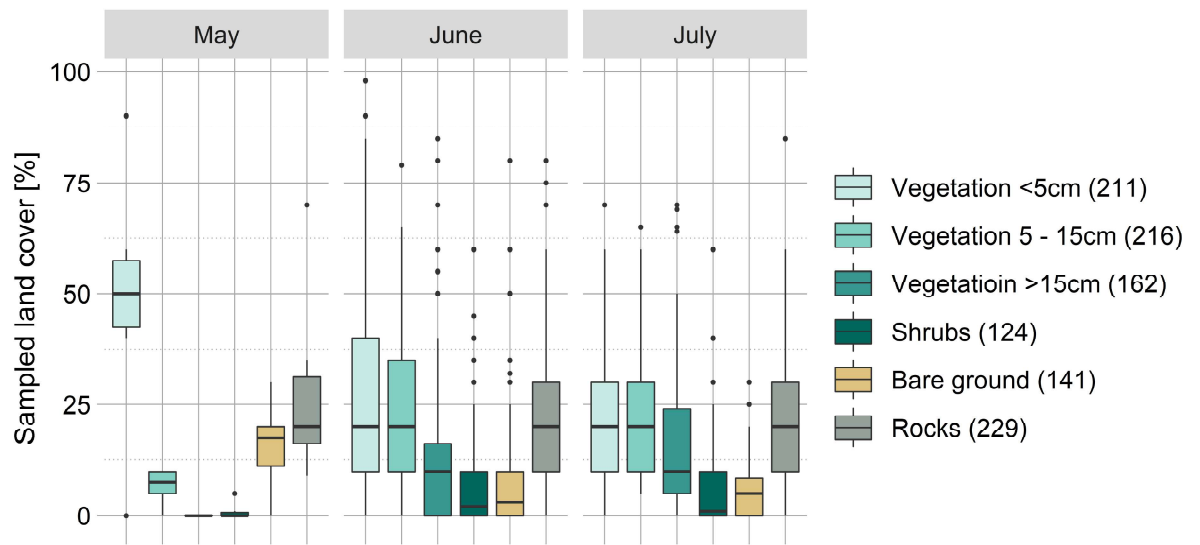
Figure 50: The percentages of vegetation <5cm and bare ground decreased during the course of the year, whereas the coverage of vegetation 5 - 15cm and vegetation >15cm increased. The ratio of shrubs and rocks remained similar from May to July.
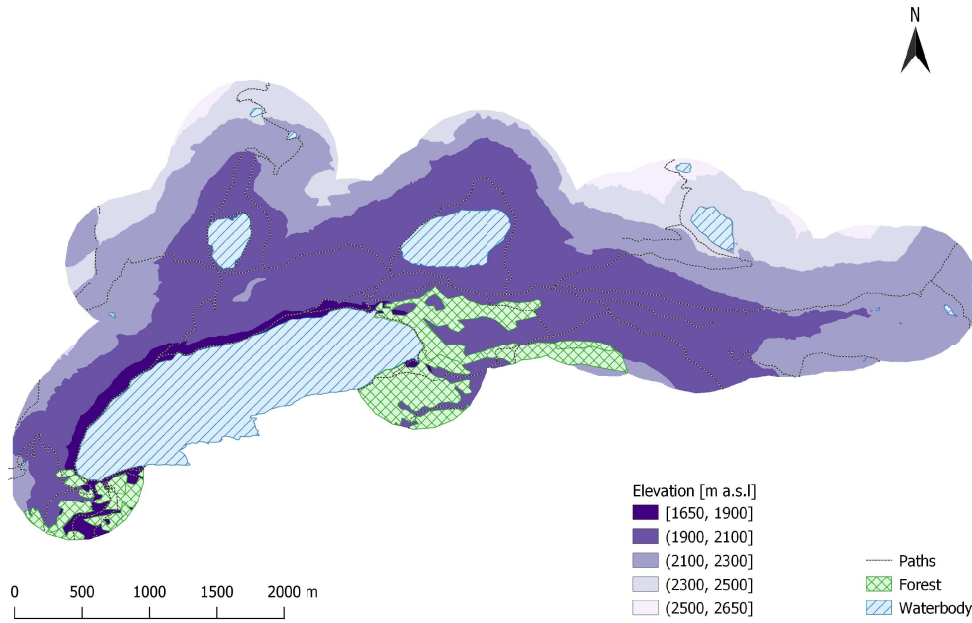
6.2.3 Digital elevation model (DEM)

Figure 51: The original elevation values in the study area depict a spatial distribution similar from the median values. Most of the study site remains below 2100m a.s.l..
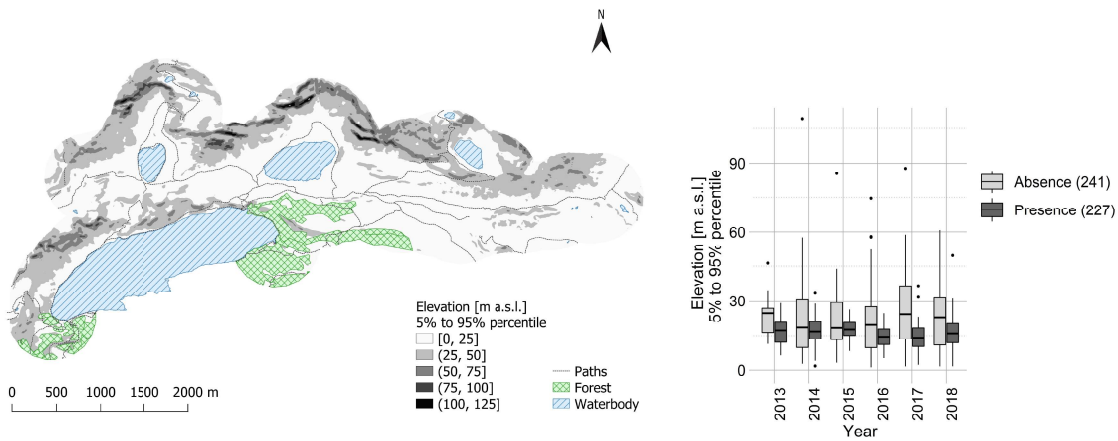
(a) The evenness of the central study area contrasts the steep cliffs in the northern part.

(b) The elevation of the presence points varies less.

Figure 52: The 90% percentile distance of elevation reports the same characteristics as median of slope, see Figure 7.

6.2.4 Slope

The original slope value reports small-scaled variation, which disappeared in the median, see Figure 7. However, the 90% percentile distance reveals these deviations, see Figure 8. Therefore, the combination of slope median and slope 90% percentile distance approximately describes the spatial distribution of the original values of the slope.
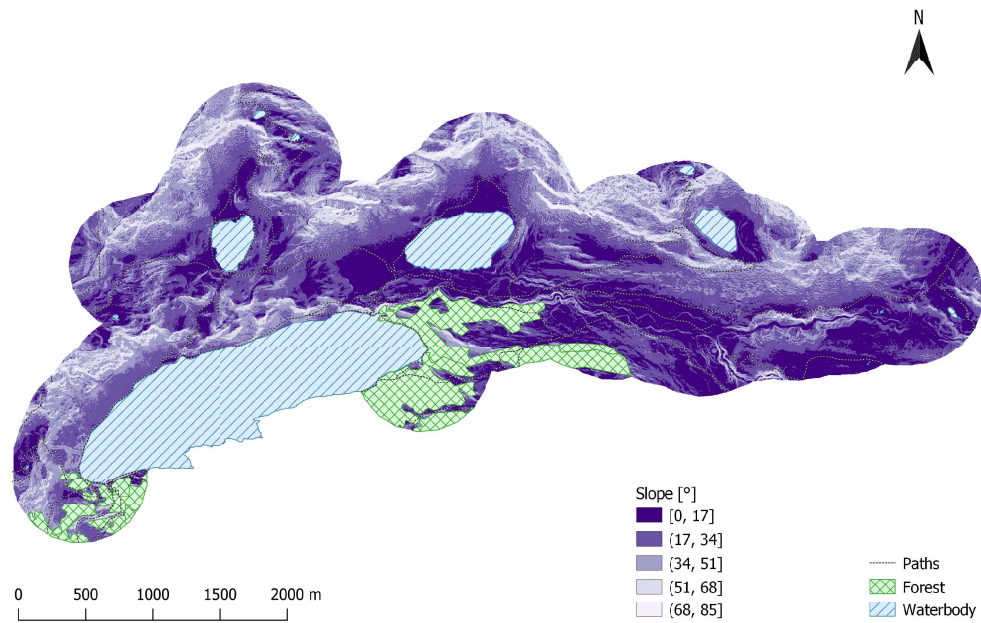

Figure 53: The Lago Cadagno lies in an even and flat marsh, whereas the Murinascia Grande flows through a mostly even but slightly varying plain between the Lago di Dentro and the Lago Ritom. The steepest parts are the gorge of the Murinascia Grande and the cliffs in the northern part of the study area and at the northern shore of the Lago Ritom.
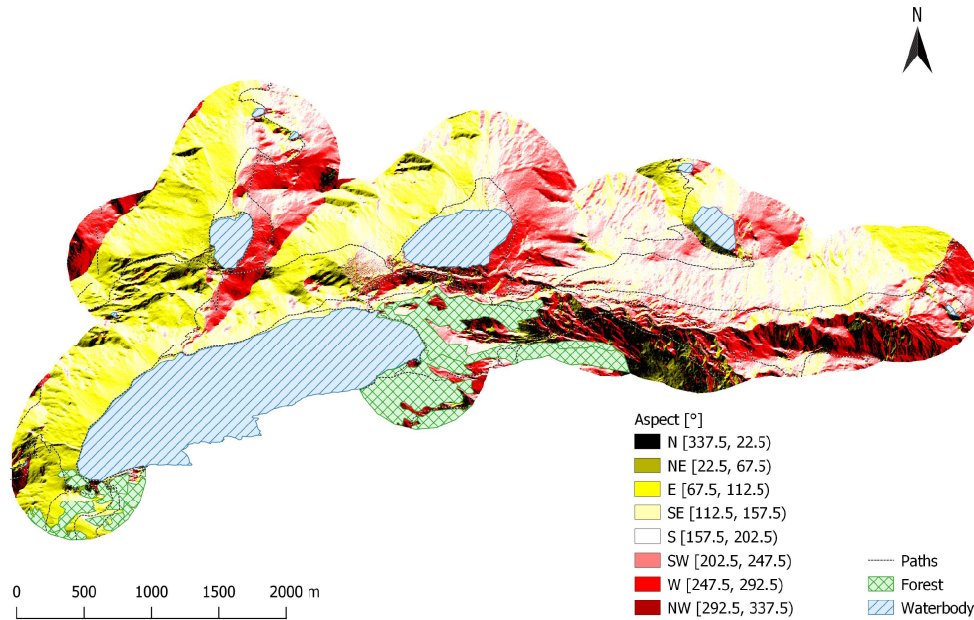
6.2.5 Aspect

Figure 54: The original values of aspect depict similarly to the median values an east-west orientation of the eastern part and a north-south axis through the western part of the study area. The gorges western of the Lago di Tom emerge more distinctively than in the median values, see Figure 9.

6.2.6 Roughness
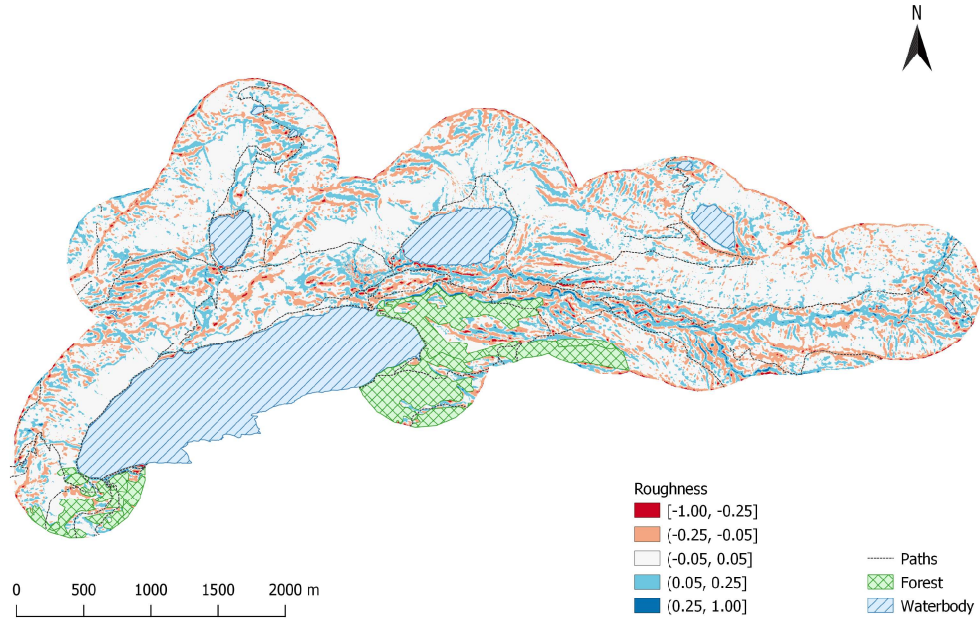
Figure 55: The original values of roughness are more particularised than the median values. Yet, its coarse granularity prevents the identification of small-scaled features such as cairns or the alternating narrow gorges and ridges of alluvial fans.
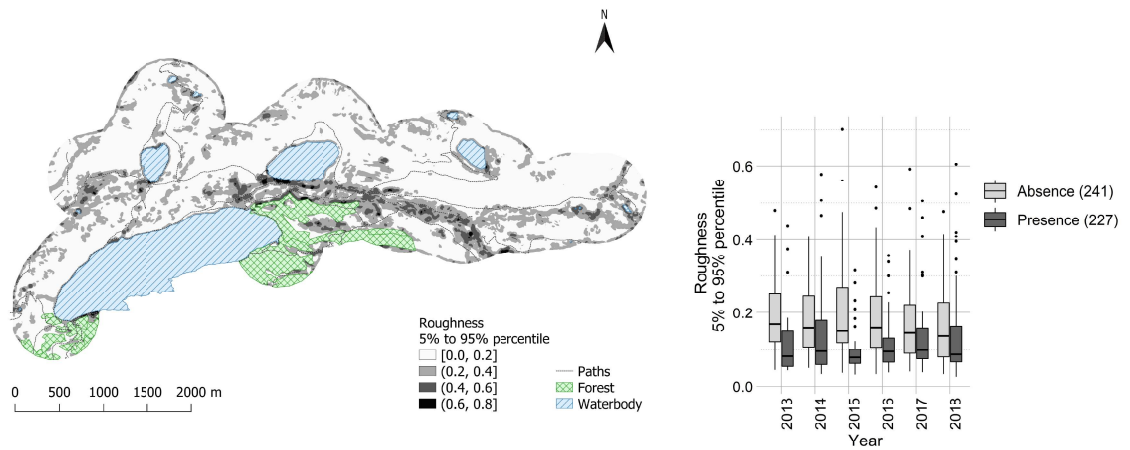
(a) The largest variation in roughness occur in the gorges of the Murinascia Grande, western of Lago di Tom and at the bottom of the steep cliffs.

(b) The presence points range less in the roughness than the absence points.

Figure 56: The 90% percentile distance of the roughness highlights similar elements as the 90% percentile distance of the aspect.

6.2.7 Normalized difference vegetation index (NDVI)
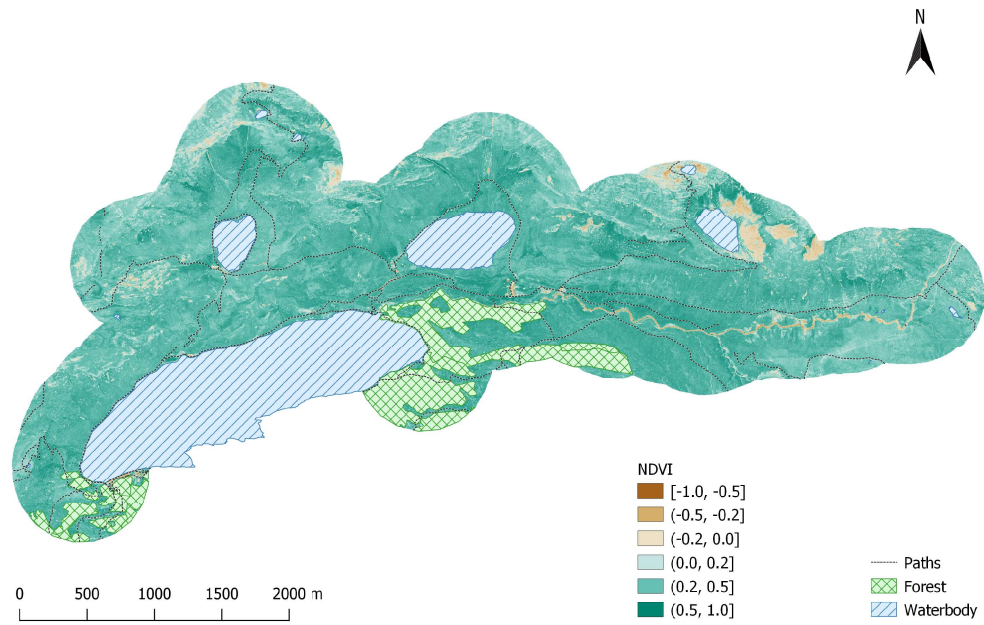
Figure 57: Only the original values of the NDVI disclose the multitude of cairns in the pastures, e.g. at the bottom of the cliffs between the Lago di Dentro and the Lago Cadagno. The NDVI successfully differentiates unvegetated from vegetated surfaces.

6.2.8 Number of temporary snow-free days
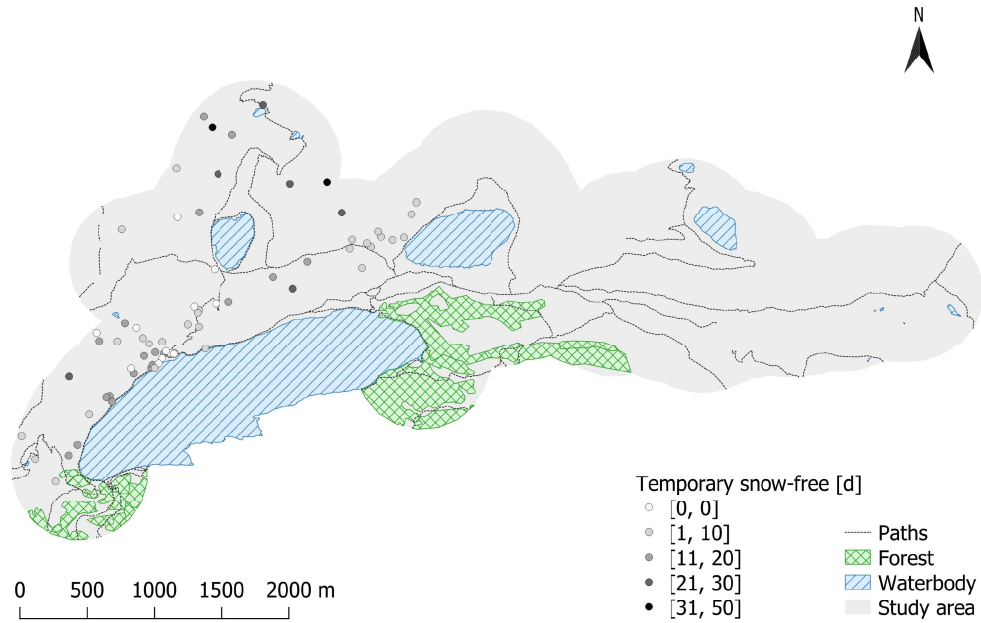
Figure 58: The number of temporary snow-free days equals the number of snow-free days between the first snow-free day and the last snow-covered day. Few temporary snow-free days appear mostly at the northern shore of the Lago di Ritom and in the west of the Lago Cadagno. Many temporary snow-free days occur around the Lago di Tom.
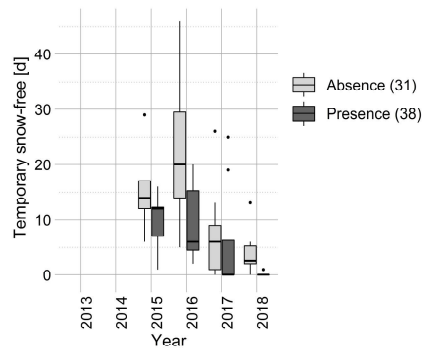
Figure 59: Presence points disclose less temporary snow-free than the absence points of the same year. There is a large between-year variation.

6.2.9 Cavities
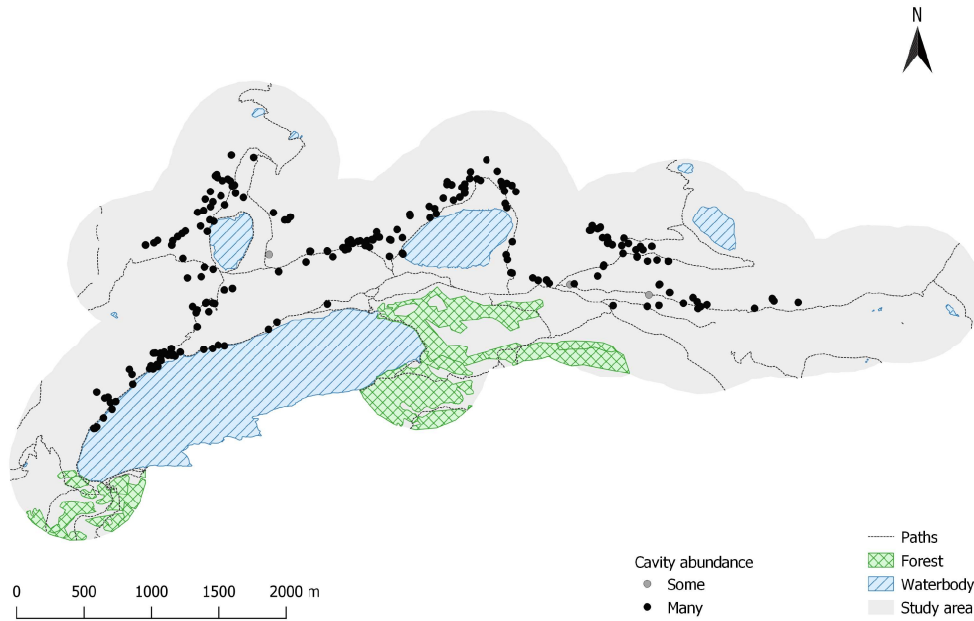
Figure 60: The nest is usually located in cavities in walls, cairns, crevices, marmot burrows and other burrows [Corti and Melcher 1958; Glutz von Blotzheim 1962; Glutz von Blotzheim and Bauer 1988; Wartmann 1985]. The abundance of cavities is uniformly high throughout the study area.

6.2.10 Perches

Perches overlook the surroundings and benefits the northern wheatear not only as song posts but also as starting point for hunting. Perches in the Val Piora mostly consisted of rocks, trees, cattle fence poles and shrubs. I assessed their height approximative using the categories up to knee high (<0.5m), up to hip (0.5 - 1.0m) and hip and higher (>1m), see Figure 62 and 63.
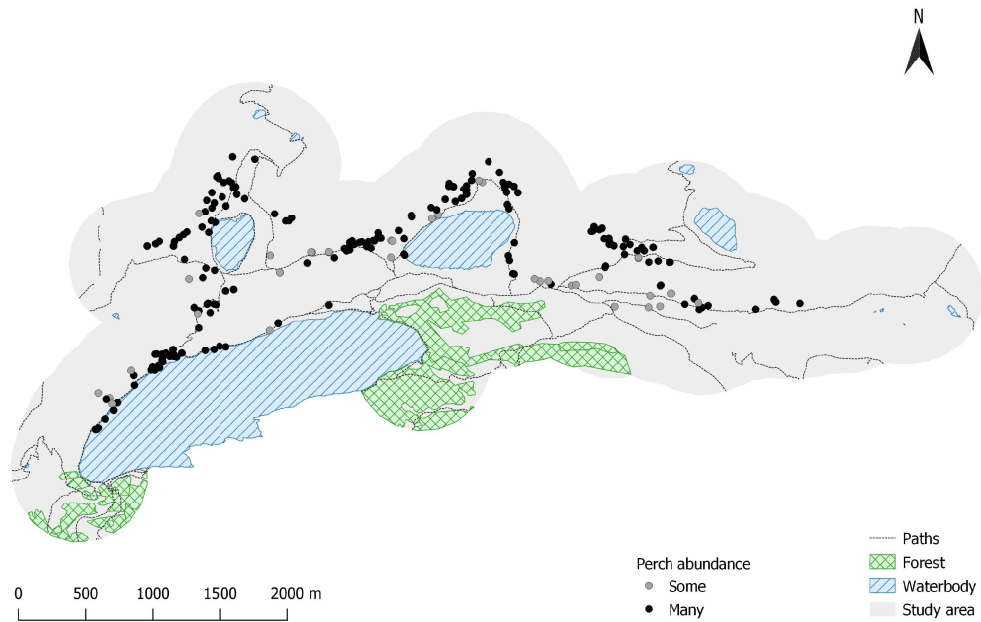
Figure 61: Perches are almost as omnipresent as cavities in the study area.
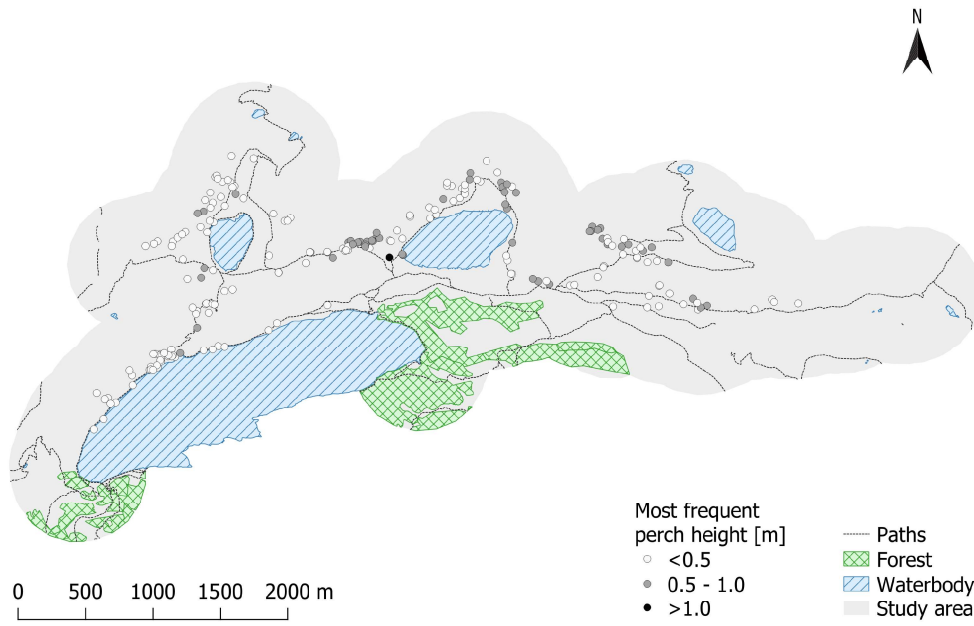
Figure 62: The most frequent found perches were rocks with heights <0.5m. Other often found perches were cairns and cattle fence poles, both hardly ever exceeded 1m.
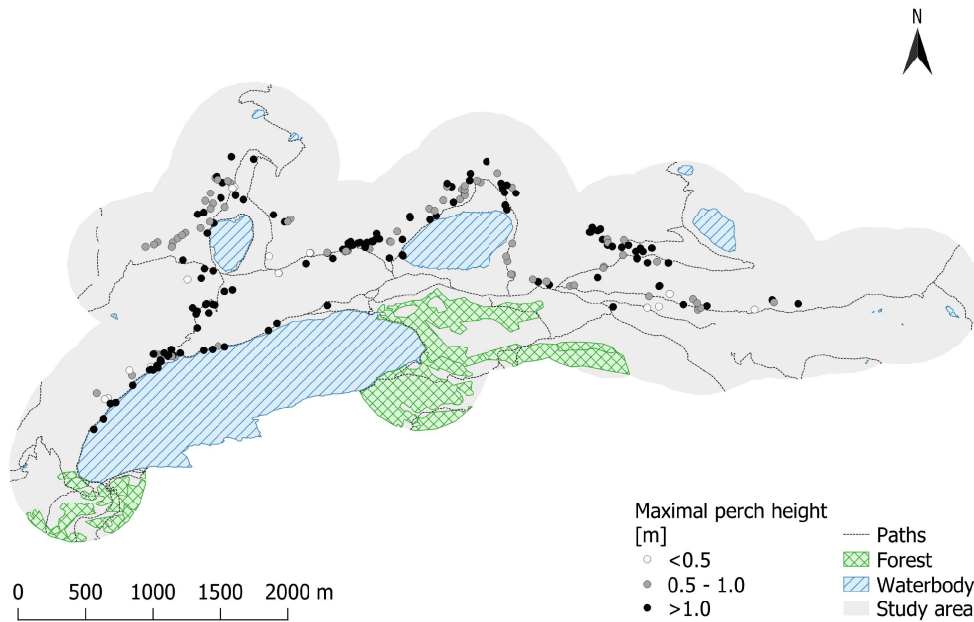
Figure 63: The highest observed perch of a nest site were mostly enormous rocks, cliffs or stone walls. These were generally higher than one metre.

6.2.11 Cairns
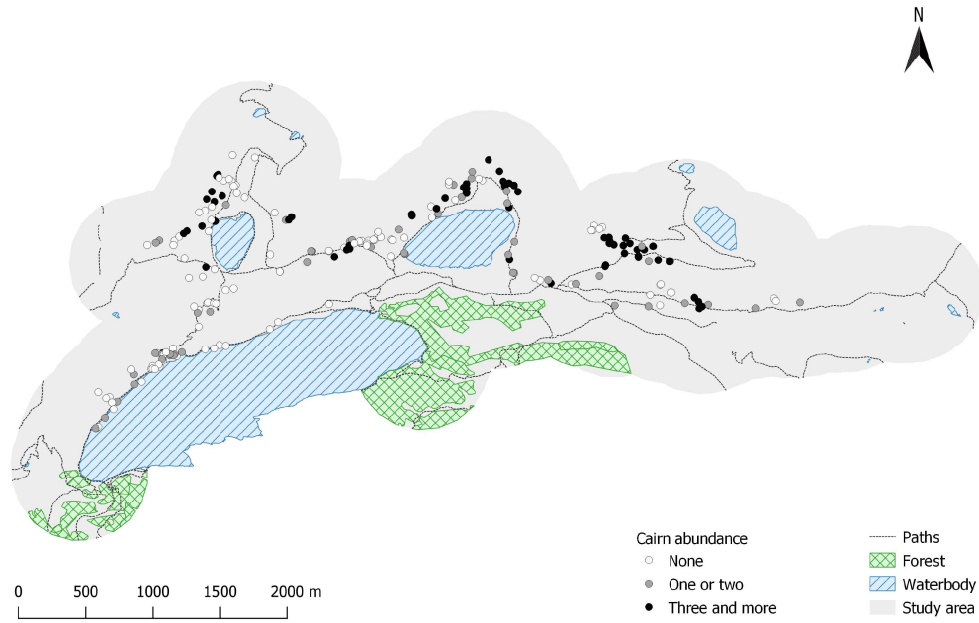
Figure 64: A cairn is a pile of stone created from stones found in the surrounding meadows. Consequently, they were most abundant at the bottom of steep cliffs where fallen rocks scattered the meadows.

6.2.12 Pairs plot

A pairs plot compares a set of variables pairwise. The name of a variable is written over its histogram in the diagonal. Pairwise correlation coefficients are below the diagonal; their font size represents the magnitude of the correlation. Pairwise scatter plots, including a fitted function, are above the diagonal. Variables with a high correlation (e.g. >0.6, see DEM_perc and slope_med in Figure 66) depict a visible trend in the scatter plots and are advisably excluded from statistical models [Meier 2018]. The numbers on the top of the pairsplot are in brackets in the legend and replace the unreadable variable names.
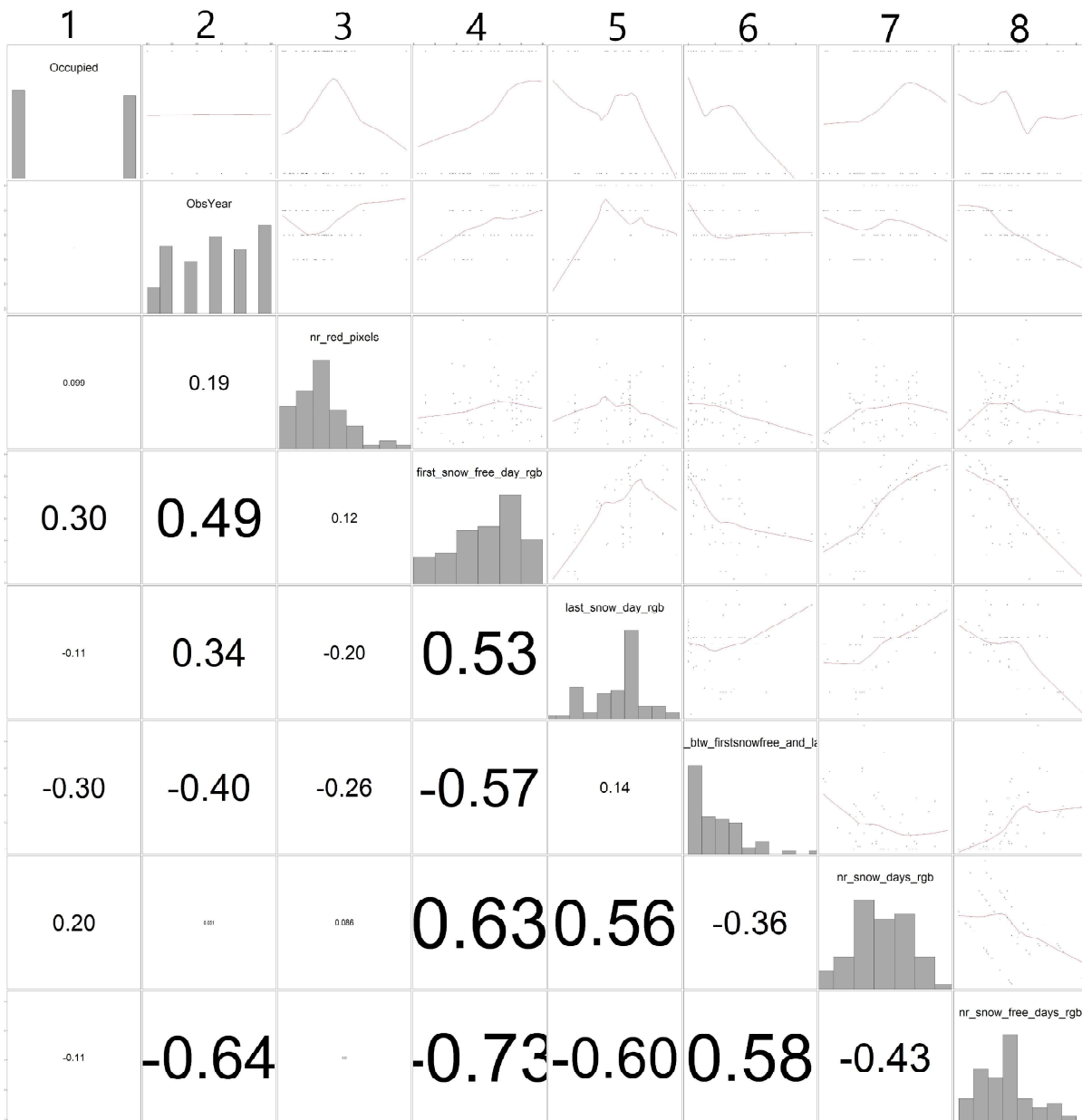
Figure 65: This pairs plot contains 68 data points with the computed snow coverage values. (1) Occupied, (2) observation year, (3) size of the nest site [pixel], (4) first snow-free day, (5) last snow-covered day, (6) number temporary snow-free days, (7) total number snow-covered days and (8) total number snow-free days. There is no strong correlation between the size of the nest site and the found snow coverage, see the third column. A correlation was found for first day snow-free, last day snow-covered respectively temporary snow-free with number of days snow-covered and number of days snow-free.
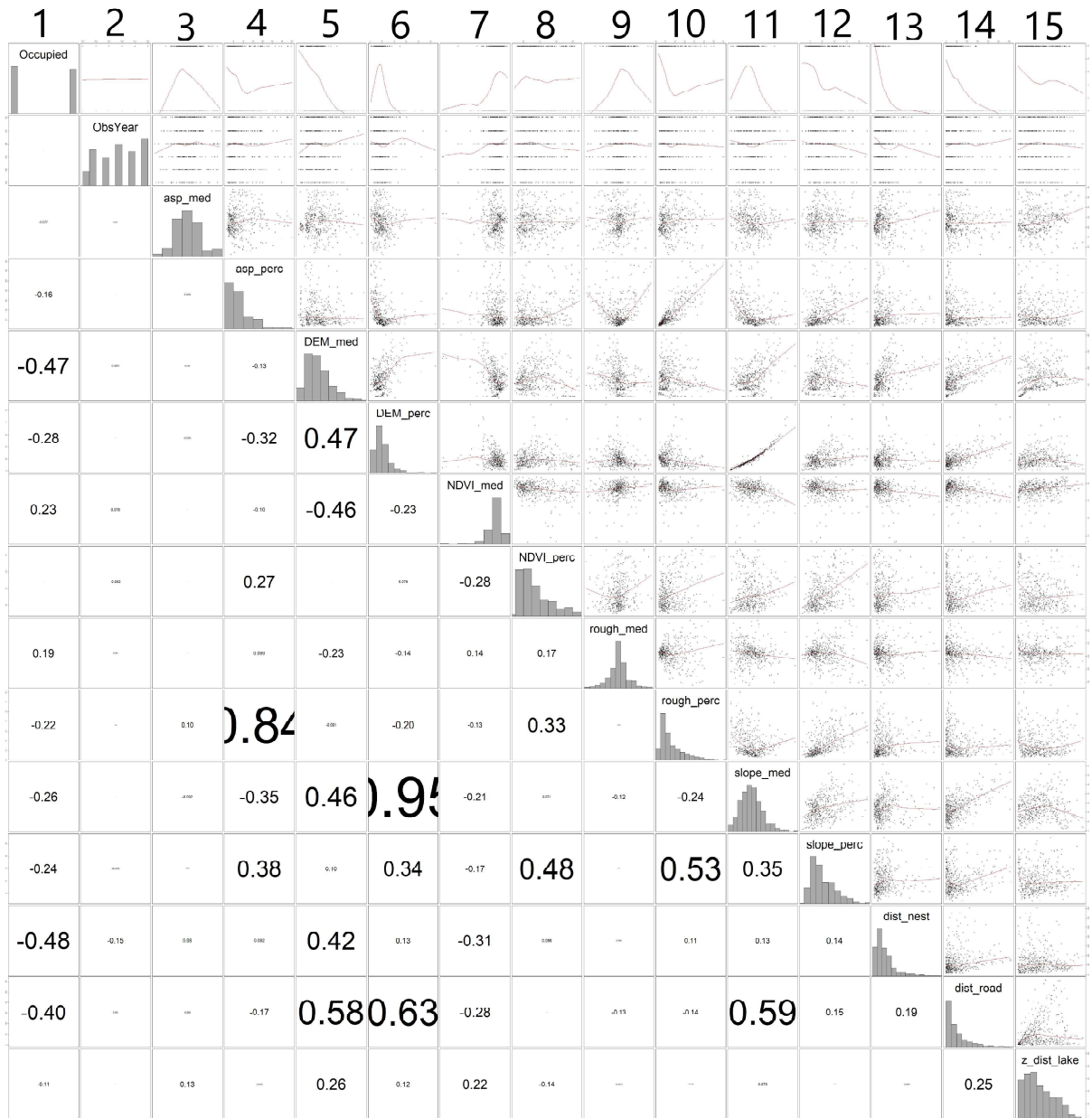
Figure 66: This pairs plot contains 468 data points. These presence and absence points were used in the GLM without snow. (1) Occupied (presence or absencee point), (3) median of aspect, (4) 90% percentile distance of aspect, (5) median of elevation, (7) median of NDVI, (8) 90% percentile distance of NDVI, (9) median of roughness, (11) median of slope, (12) 90% percentile distance of slope, (13) cumulated distance to the three nearest nests and (14) distance to the nearest path. The variable (15) distance to the nearest lake was only used to exclude points situated on the shores of the Lago Ritom. The (2) observation year was excluded from the model as it caused fitting difficulties of the maximal GLM while the maximal GLMM without snow showed no effect of this variable. The variable (6) 90% percentile distance of elevation, respectively (10) 90% percentile distance of roughness, was excluded as it shows a high correlations with slope median, respectively aspect 90% percentile distance.
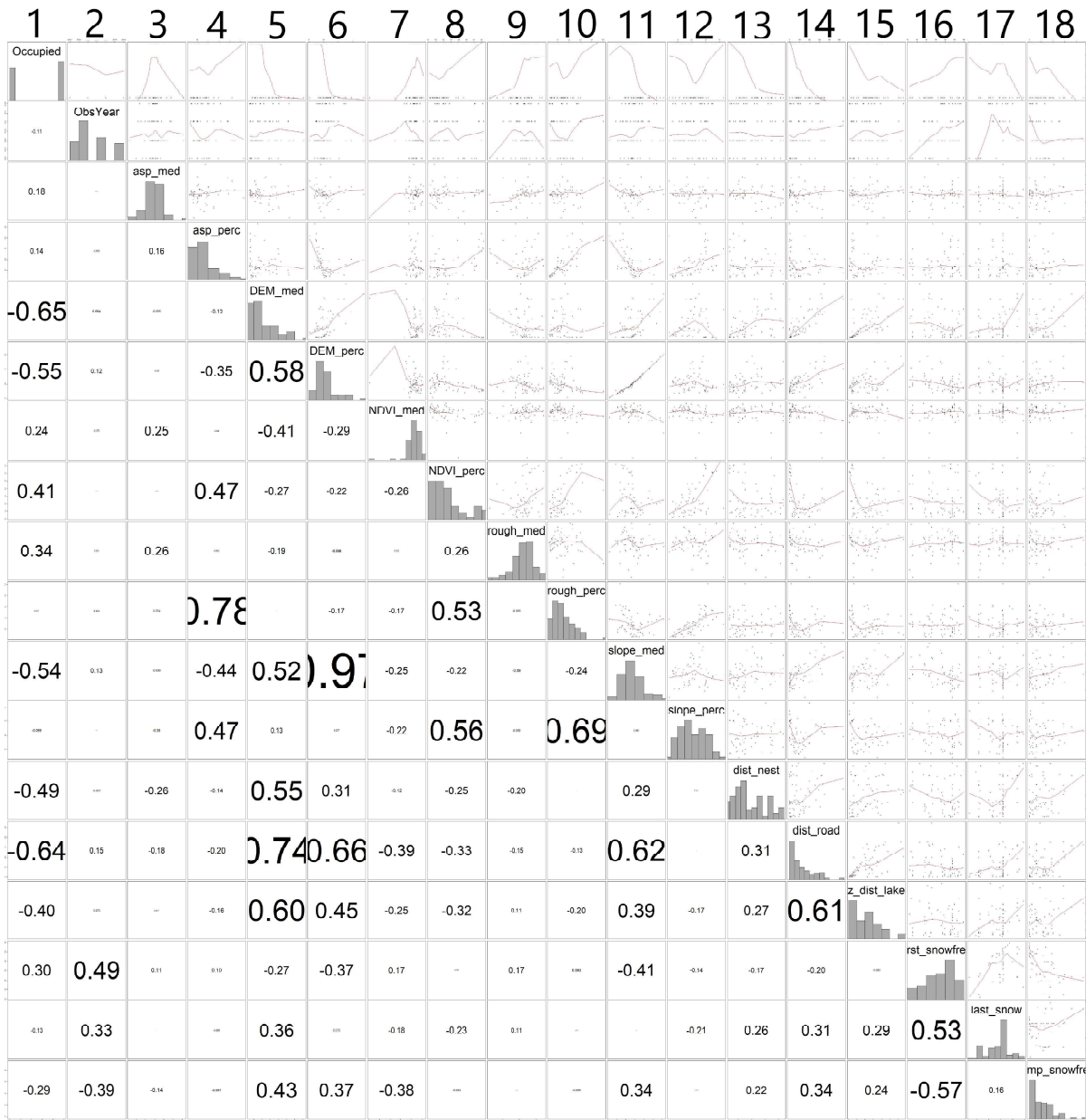
Figure 67: This pairs plot contains 68 data points for which the snow coverage was analysed. These presence and absence points were used in the GLM with snow. (1) Occupied, (3) median of aspect, (5) median of elevation, (7) median of NDVI, (11) median of slope, (9) median of roughness, (13) cumulated distance to the three nearest nests, (16) first snow-free day and (17) last snow-covered day. Because of model fitting difficulties, I excluded the variables (2) observation year, (4) 90% percentile distance of aspect, (6) 90% percentile distance of elevation, (8) 90% percentile distance of NDVI, (10) 90% percentile distance of roughness, (12) 90% percentile distance of slope, (14) distance to the nearest path and (18) temporary snow-free. (15) Distance to the nearest lake was used to exclude the nests on the shore of the Lago Ritom.

6.2.13 RGB threshold assessment

The snow coverage uses a RGB threshold value above which a pixel is considered snow. I determined this threshold empirically by sampling a number of *Photos* taken at the Lago di Tom in 2017. I considered snow and similarly looking surfaces such as concrete, gravel and white water. The examined spectral properties were the minimum (min) of the red (R) band and its mean, the minimum of the green (G) band and its mean, the minimum of the blue (B) band and its mean. Adobe Photoshop provided these properties for patches of pixels.

image	sample	R min	R mean	G min	G mean	B min	B mean	image description
Tom1700004	white snow	250	254	250	255	245	255	snow, sunny
Tom1700004	shadowy snow	205	215	205	220	225	239	snow, sunny
Tom1700004	shadowy snow in forest	155	168	150	170	220	239	snow, sunny
Tom1701472	snow with shadow spots	170	206	165	206	210	240	snow, shadowy
Tom1701472	bright snow	205	218	210	221	230	239	snow, shadowy
Tom1701472	shadowy snow	180	195	180	196	230	243	snow, shadowy
Tom1701472	shadowy snow	145	178	140	178	220	250	snow, shadowy
Tom1701493	snow/ice	140	192	155	213	190	245	snowmelt, blue, shadowy
Tom1701493	shadowy snow	165	200	200	225	220	245	snowmelt, blue, shadowy
Tom1701493	snowmelt with dirt	75	175	80	198	175	245	snowmelt, blue, shadowy
Tom1701511	snow/ice	180	226	175	228	200	242	snowmelt, shadowy
Tom1701511	snow/ice	175	220	170	222	215	243	snowmelt, shadowy
Tom1701511	snow with dirt	210	240	215	244	220	245	snowmelt, shadowy
Tom1701511	dirty snow	125	225	100	225	185	240	snowmelt, shadowy
Tom1701511	dirty snow	150	220	150	220	190	240	snowmelt, shadowy
Tom1701511	dirty snow	145	217	150	220	200	240	snowmelt, shadowy

Tom1701524	snow/ice	250	254	245	254	245	254	snowmelt, sunny
Tom1701524	snow	245	253	245	252	245	253	snowmelt, sunny
Tom1701524	melting snow	205	236	190	229	210	237	snowmelt, sunny
Tom1701524	melting snow	210	239	200	233	210	240	snowmelt, sunny
Tom1701532	snow/ice	145	180	145	185	220	246	remaining snow, shadowy
Tom1701532	dirty snow	150	200	145	209	200	242	remaining snow, shadowy
Tom1701532	snowfield	225	245	230	250	220	245	remaining snow, shadowy
Tom1701532	shadowy snow	130	188	135	192	215	247	remaining snow, shadowy
Tom1701532	shadowy snow	160	194	155	195	195	244	remaining snow, shadowy
Tom1701550	snow/ice	170	213	170	215	200	240	remaining snow
Tom1701550	snow	245	252	247	252	245	252	remaining snow
Tom1701550	dirty snow	195	234	185	237	205	244	remaining snow
Tom1701555	snow/ice	250	254	250	254	250	254	remaining snow, sunny
Tom1701555	snowfield	200	252	200	251	200	252	remaining snow, sunny
Tom1701555	melting snow	215	240	205	235	205	232	remaining snow, sunny
Tom1701832	ice	155	168	150	162	210	220	new snow
Tom1701832	snow	190	234	190	233	190	240	new snow
Tom1701832	shadowy snow	185	205	185	205	225	240	new snow
Tom1701832	dirty snow	210	232	205	230	220	241	new snow
Tom1701832	shadowy snow	180	195	180	198	225	241	new snow
Tom1703204	snow	205	225	195	224	225	244	spring
Tom1703232	snow	245	251	245	251	245	250	spring, sunny
Tom1703204	whitewater	230	252	235	254	230	249	spring

Tom1703204	concrete	130	173	137	169	160	207	spring
Tom1703232	bright beach	237	250	239	250	240	249	spring, sunny
Tom1703232	concrete	205	244	190	237	216	240	spring, sunny

Table 10: The results of these empirically determined values suggest that I could not detect snow when (1) the image is bluish, (2) the snow lies in shadow, (3) the snow is dirty. On the other hand, I would detect snow (falsely) if (1) concrete, (2) gravel beaches, (3) hiking trails or (4) water were analysed on a sunny generally snow-free image.

Finally, the used threshold was empirically determined to be ~155 for red and green and ~190 for blue. I chose the RGB values of snow on *Photos* where snow was not bright white yet still spectrally distinguishable from the other surfaces. Thus, I aimed on a threshold able to deal with boundary cases. The slightly increased value for the blue band originated in the snow showing a slightly bluish cast.

6.2.14 Cameras

Each season, three cameras recorded parts of the study area from three locations, see Table 11. Two different kinds of cameras recorded time-lapse videos. HYPERFIRE by RECONYX took an image at 9AM and 3PM each day and GardenWatchCam by brinno took an image every hour. I restricted the snow coverage analysis to the *Photos* taken by the GardenWatchCam because this camera recorded most of the footages. However it corrected less for different illumination conditions.

Lake	X	Y
Lago di Tom	2695576	1154436
Lago Cadagno	2696768	1154499
Lago Ritom	2696821	1154588

Table 11: Camera positions per lake in LV95.

The GardenWatchCam was used at the Lago di Tom in 2014 and at the Lago Ritom, Lago Cadagno and Lago di Tom in 2015, 2016, 2017 and 2018. The HYPERFIRE recorded the Lago Ritom, Lago Cadagno and Lago di Tom in 2013 and Lago Cadagno and Lago Ritom in 2014. The cameras at the Lago Cadagno and especially at the Lago di Tom wiggled strongest. Thus the recorded area changed over the course of the year or even within one day. Table 12 discloses qualitatively the wiggling of the camera.

Year	Lago Ritom	Lago Cadagno	Lago di Tom
2013	not analysed	not analysed	not analysed
2014	not analysed	not analysed	mostly stable
2015	stable	very shaky	mostly stable
2016	stable	mostly stable	mostly stable
2017	stable	very shaky	mostly stable
2018	stable	very shaky	very shaky

Table 12: The stability of the cameras varies. While the camera at the Lago Ritom usually recorded a stable scene, the wiggling of the camera at the Lago di Tom varied and the camera at the Lago Cadagno generally wiggled.

6.2.15 Prediction map

Originally, I intended to produce a nest suitability map for the entire study area. I planned to display its statistical certainty on a second map. Its practical accuracy could have been tested the following summer season in the study site, e.g. by searching more thoroughly for nests in the areas marked suitable on the map. No found nests in apparently suitable areas or nests in assumedly unsuitable areas would indicate that the model was lacking an important factor. The model was:

$$glmer(Occupation \sim aspect.median.z+$$
$$distance\ to\ the\ nearest\ path+$$
$$(1|ObsYear), \tag{16}$$
$$data = dat,\ family = binomial)$$

I assessed the model assumptions as in 2.5.2 and investigated the effects of 10'000 simulations of the found model.

I empirically determined the threshold between presence and absence points. The boxplots of the presence and absence points revealed that a prediction index of 0.63 discriminates the two boxes, see Figure 68. Each of this box spans between the 25% and 75% percentile. I decided to use 0.63 rather than 0.5 as threshold value as it seemed more data specific.
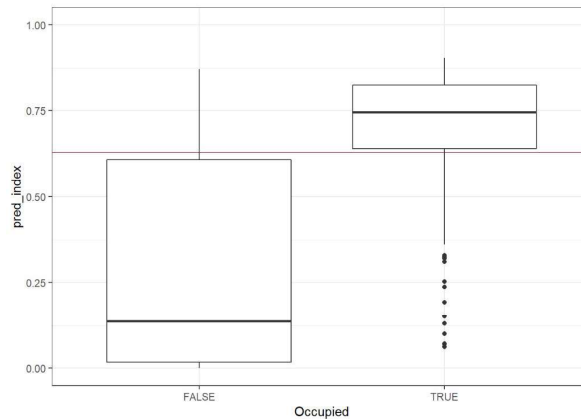
Figure 68: The pred_index describes the found index of occupancy, with zero indicating absence and one indicating presence. Presence and absence points seem to be separated at the 0.63 pred_index.

Then, I mapped for each point the found occupancy and the modelled occupancy, see Figure 69 [F. Korner-Nievergelt and P. Korner-Nievergelt 2019]. Latter originated in one randomly drawn prediction of the computed model and was calculated only for the sampled points.
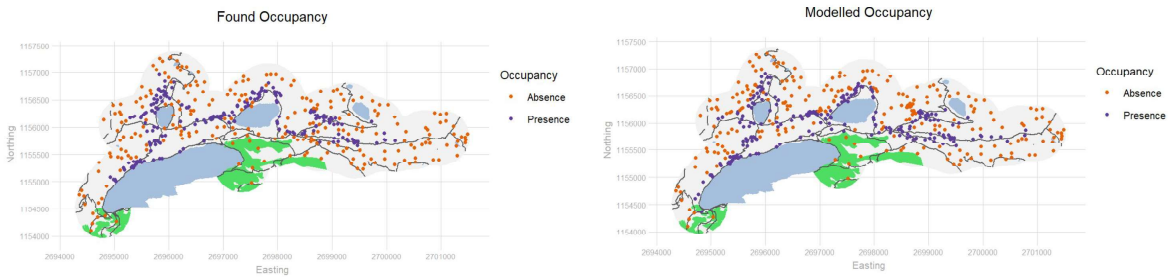
(a) This is the found spatial distribution of presence and absence points. (b) This is the predicted spatial distribution of the presence and absence points.

Figure 69: On first glance, the found and predicted distributions are similar.

The differences or agreements of prediction of each point in Figures 69a and 69b is shown in 70. The following map spares the visual point-wise comparison and aids in the detection of spatial patterns.
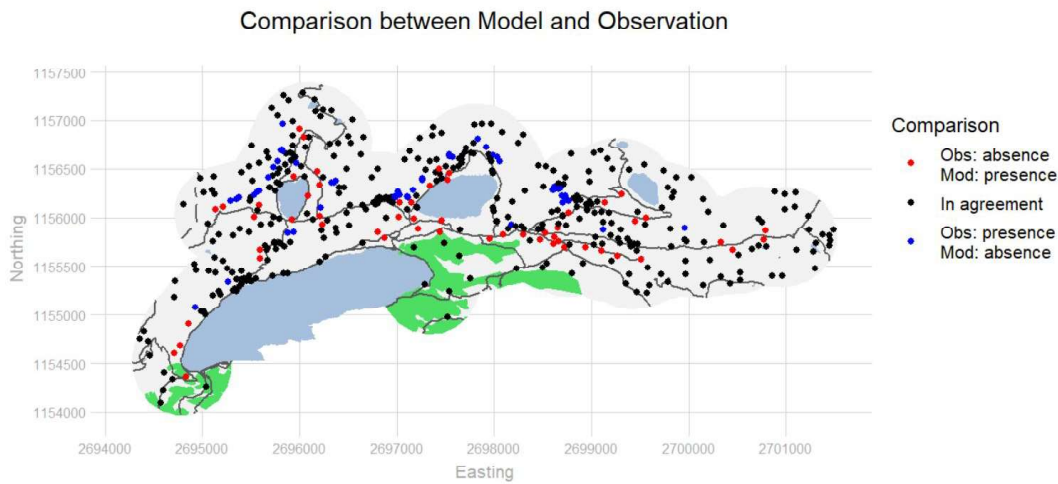
Figure 70: Black points signify agreement between the found and modelled occupancy. Red and blue indicate disagreement.

The found basic model already predicted the majority of points correctly. Most incorrectly predicted points were located around the Lago di Tom, Lago Cadagno and the Capanna Cadagno. The predicted occupancy is currently derived from a single simulated data set. However, the mean occupancy of at least 10000 simulated data set would provide a more robust map [F. Korner-Nievergelt and P. Korner-Nievergelt 2019]. Furthermore, I needed to implement a measure of certainty [F. Korner-Nievergelt and P. Korner-Nievergelt 2019]. Obviously, the colour scheme and legend require refinement.

To analyse the entire study area, an area-covering occupancy index was needed. I wished to try the coding of a predicted raster in Figure 71.
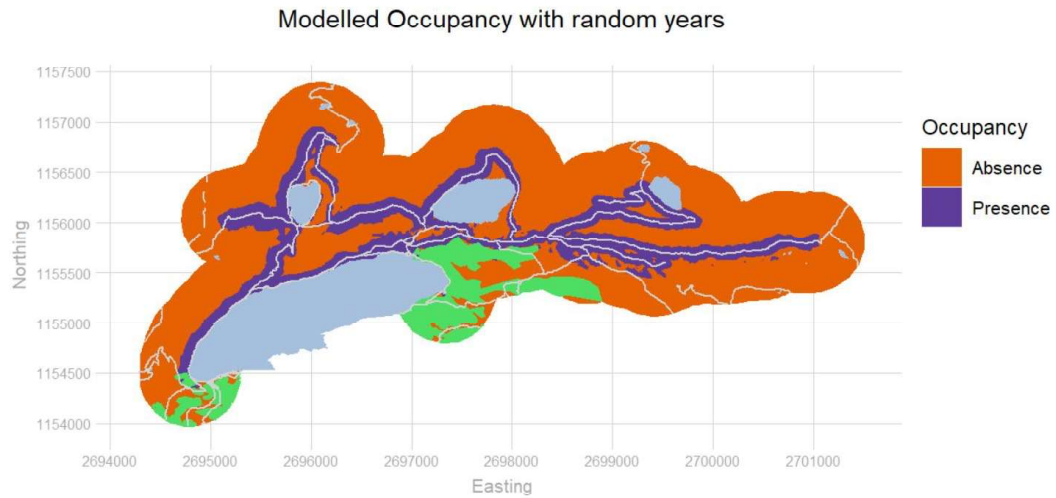
Figure 71: This map predicted the occupancy index for the entire study area. The reduced model of distance to the nearest path and median of aspect predicted presence points around hiking paths.

I showed that I am able to predict a new raster by predicting a model on an existing set of rasters. This model seems to indicate that the distance to the nearest path stronger influences the nest site choice than the median of aspect. In Figure 71, the observation year was randomly created.

I noticed that upcoming area-covering suitability analysis could be affected by the runoff of the Murinascia Grande, the river in the eastern part of the study area. Northern wheatears do not nest in water. The snowmelt-induced runoff regime of the the Murinascia Grande peaks in May and June [Bundesamt für Umwelt BAFU 2018]. Indeed, no nests were located in the bed of the Murinascia Grande, the river in the eastern part of the study area. However, the riverbed appeared partly as bare ground instead of water on the NDVI raster. Possibly, the river was almost dried out at the end of August 2015 as that summer was poor in precipitation and particularly hot [MeteoSchweiz 2016]. The GLM without snow indicated a preference of bare ground, rocks and sparse vegetation as nesting grounds. Consequently, such a late sampling date might produce false presence points in riverbeds.

If I wished to apply this method on the found models, I further needed to take into consideration that some variables currently only exist at the location of the presence and absence points: Observation year, cumulated distance to the three nearest nest and, for the time being, the first snow-free day and last snow-covered day do not exist as area-covering rasters. The model predicting such a map needed to exclude, substitute or interpolate these variables so that they exist as rasters. Furthermore, an optimization of the colour scheme and legend is required and a measure of certainty needs to be implemented.

In this section, I tested how to create a prediction map on a basic statistical model. This section is neither complete nor scientifically sound; it was a mere test to see if I was able to predict a model on a raster. Nonetheless, I wished to document my basic idea, which remains unfinished because of shortage of time.

6.3 The additional data to results

6.3.1 Nest densities

The here presented nest densities are computed per year. The reference area equals a circle around a nest with radius 178m (~10ha). Such a reference area was computed for each known nest.

(a) 2013

(b) 2014

(c) 2015

(d) 2016

(e) 2017

(f) 2018

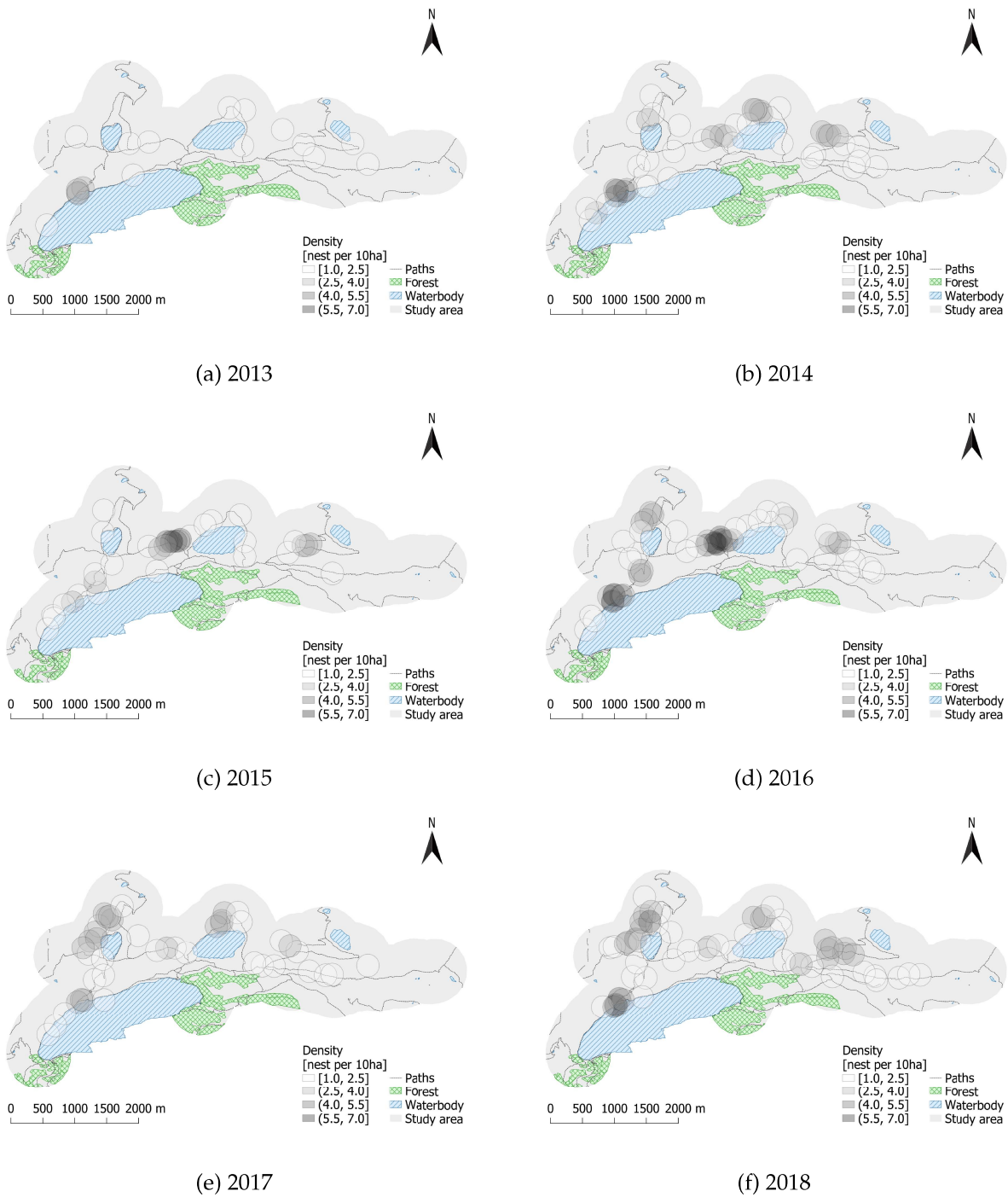Figure 72: The highest nest densities occur at the northern shore of the Lago Ritom, north-eastern and northern of the Lago Cadagno, north-eastern of the Lago di Tom and between the Lago Cadagno and the Lago di Dentro. High values per nest site signify clustering. The areas darker than the maximal legend entry origin in the stacked semi-transparent nest sites. These locations mark sites which included presumedly more than 10ha.

6.3.2 Overlay

I assessed the similarity of 15 repetitions of the overlay function of the snow coverage algorithm. To do so, I chose each time the same GCPs. The computed scaling factors and offsets in x- and y-direction are summarised in Table 13.

	scale_x	scale_y	offset_x	offset_y
1	1.00	0.88	-66.24	-186.71
2	0.99	0.87	-52.69	-181.45
3	0.99	0.87	-52.69	-181.45
4	0.97	0.87	-37.71	-181.24
5	0.95	0.88	-20.67	-180.82
6	0.97	0.87	-45.53	-180.23
7	0.98	0.88	-43.21	-188.44
8	0.97	0.87	-37.33	-182.21
9	0.96	0.87	-34.50	-182.82
10	0.96	0.87	-28.16	-182.17
11	0.97	0.87	-39.21	-178.07
12	0.96	0.87	-34.87	-177.83
13	0.97	0.87	-37.59	-182.11
14	0.98	0.87	-47.21	-178.03
15	0.98	0.87	-44.04	-180.49
Min	0.95	0.87	-66.24	-188.44
Max	1.00	0.88	-20.67	-177.83
Median	0.97	0.87	-39.21	-181.45
Range	0.05	0.01	45.57	10.61

Table 13: The minimal value, maximal value and range between 15 overlays quantify the performance of this method. The range varies strongly in x-direction while the range in y-direction is relatively narrow. The *2-Photo* was stretched if the scale factor was >1 and compressed if the factor was <1. The offset factor equalled the number of pixels by which the image was shifted up/down or left/right.

6.3.3 Snow coverage per nest site and year

Generally, the computed snow coverage data seemed to indicate three kinds of typical snow coverage courses in the Val Piora between 2015 and 2018.
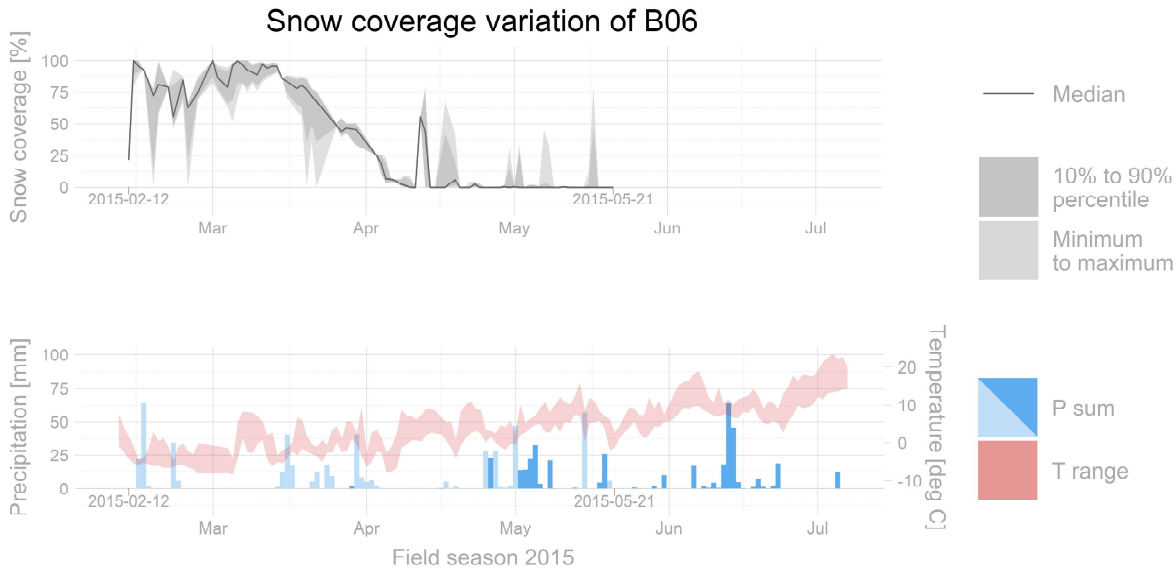
Figure 73: First, a constant snowmelt. The here disclosed snow coverage generally decreases. One minor snowfall in mid-April is visible as peak. The following peaks were smoothed by the median but not by the 90% percentile. The shown insecurity is generally low; the security depends not only on the performance of the algorithm but also on the number of analysed *2-Photos* per day.
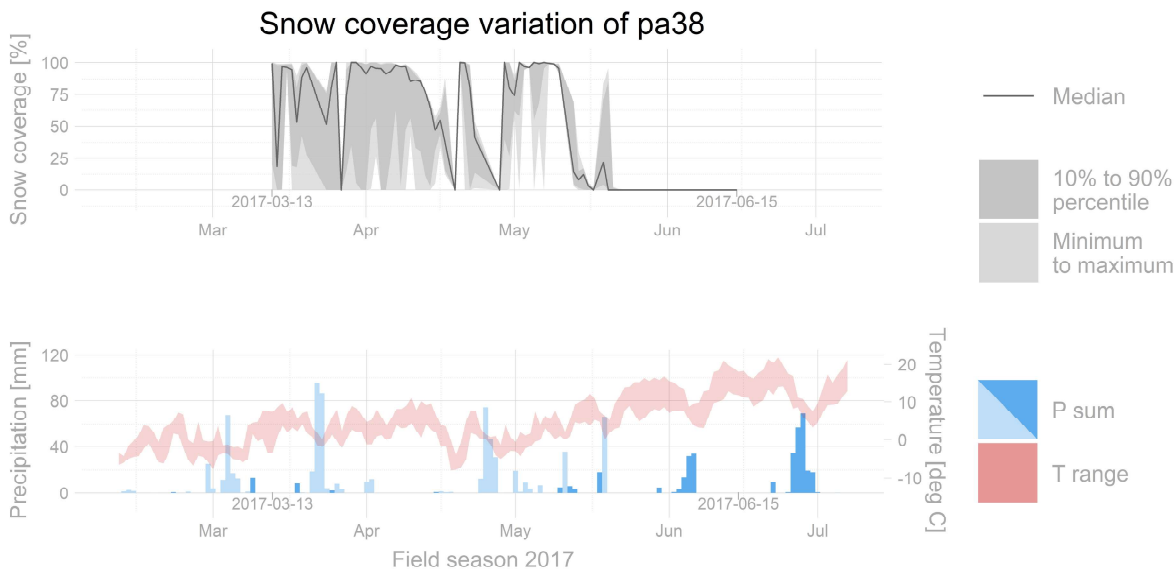
Figure 74: Second, multiple snowmelt cycles. This graph depicts three long-lasting and one light snowfall events. The general snow coverage appears smooth but with large insecurity. There is one snow-free day at the end of March. This complete snowmelt is extreme early in season. The surrounding large insecurity and high snow coverage might indicate an error in the analysis because of unfavourable illumination conditions.
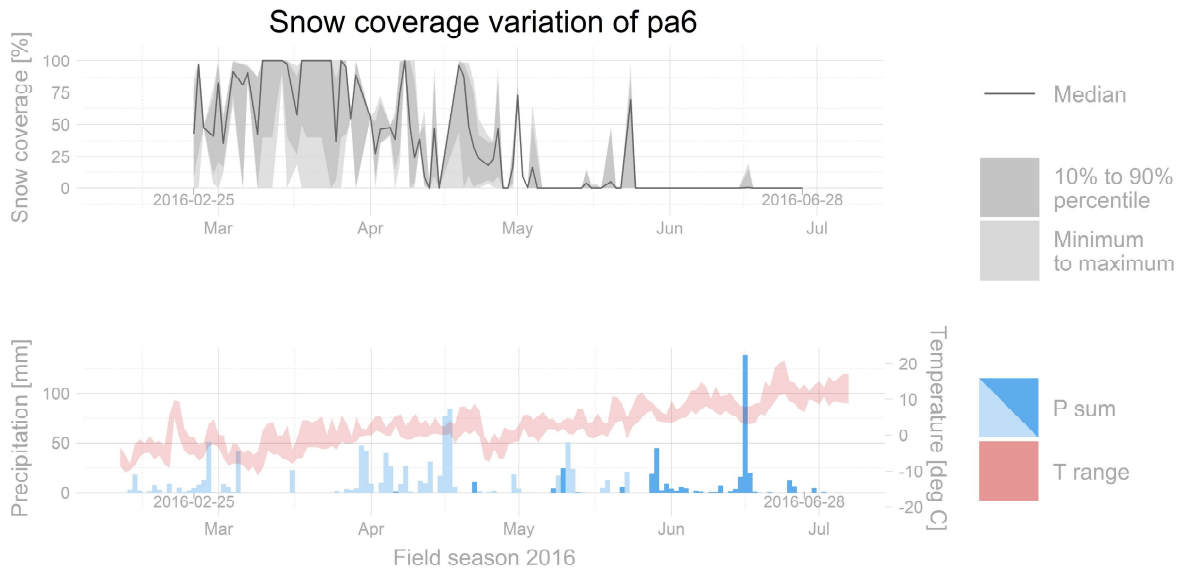
Figure 75: Despite the large fluctuations and insecurity, this plot is similar to the type one and two. There seem two major snowfall events (March to mid-April, end of April) and several minor snowfall events (start of May, end of May).
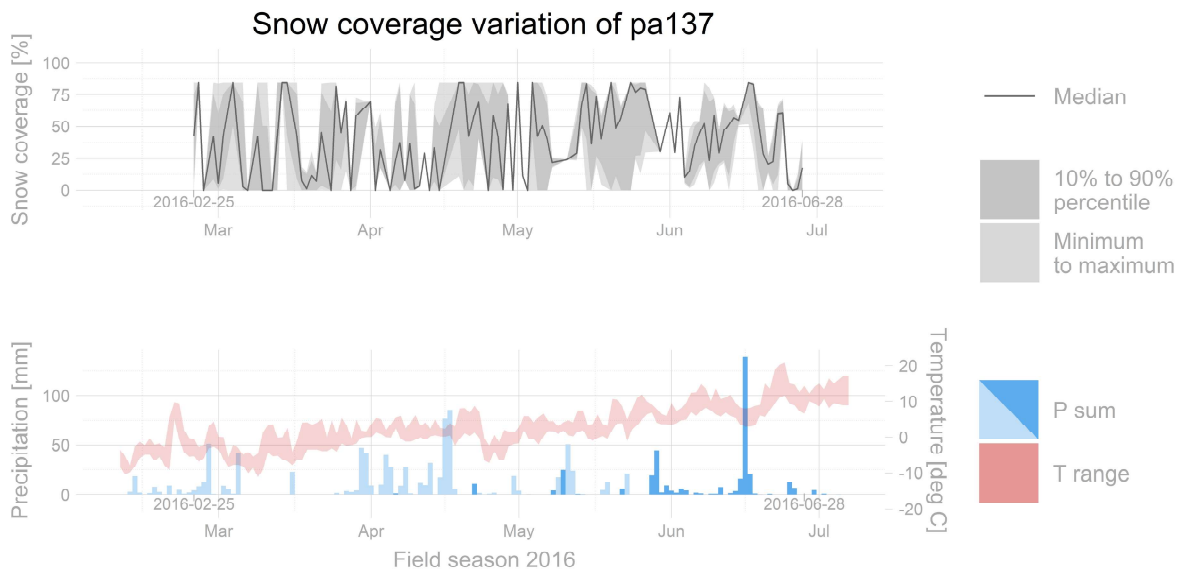
Figure 76: Third, a chaotic plot. The snow coverage zigzags wildly with a large insecurity. The true snow coverage of the nest site pa137 showed no such instability, thus the assigned quality was low.

6.3.4 Weather plots

I visualised the weather data to describe its variability in the Val Piora between 2014 and 2018. The following climatograms were the bottom plots in the output of the snow coverage algorithm.
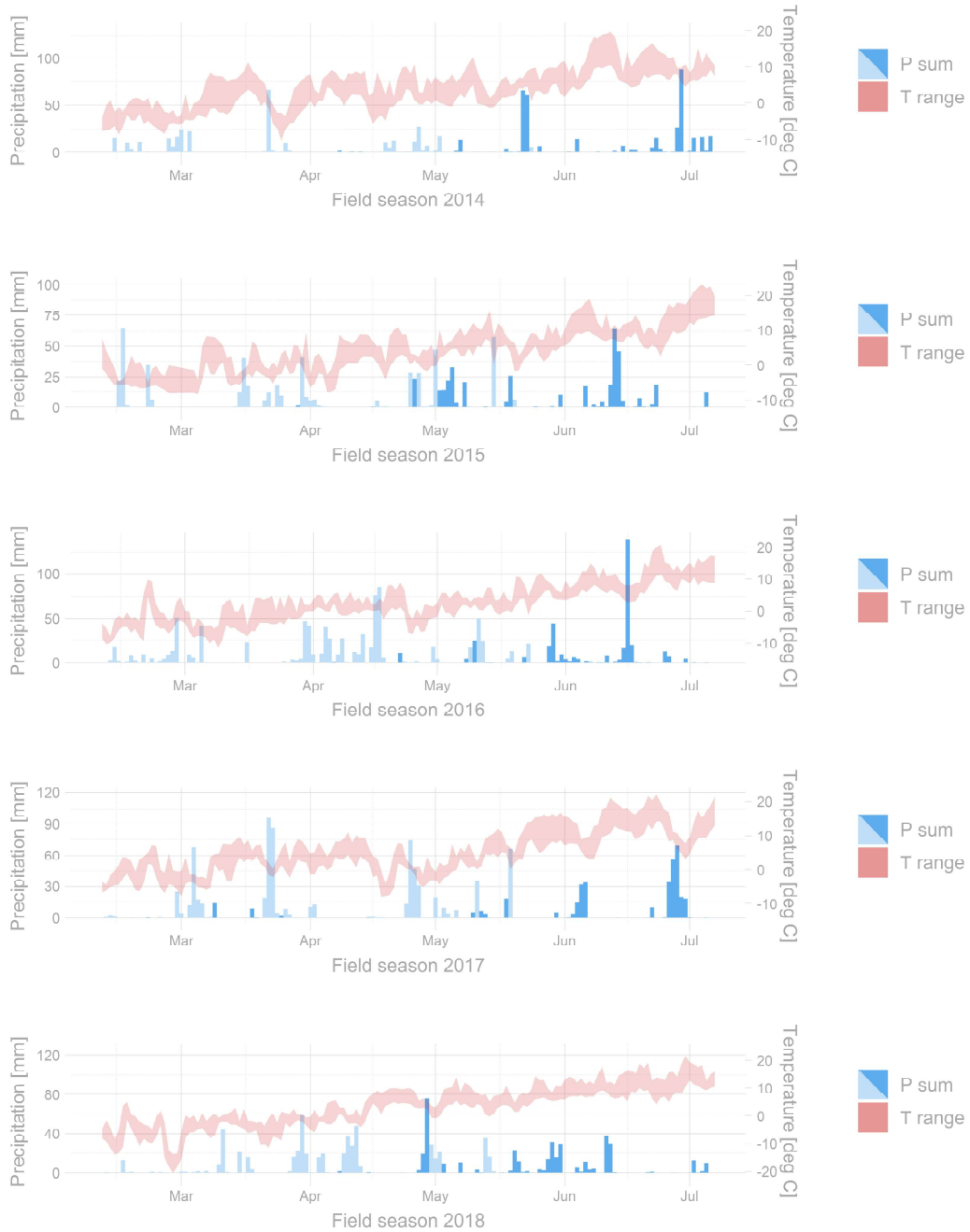
6.3.5 Linear mixed effect models (LMEM)

The model output of the first snow-free day:

```
##    Linear mixed model fit by REML ['lmerMod']
##    Formula:
##    first_snow_free_day_rgb   NDVI_median.l + NDVI_perc.l + DEM_median.l +
##    aspect_median.l + aspect_perc.l + slope_median.l + roughness_median.l +
##    NDVI_median.q + NDVI_perc.q + DEM_median.q + aspect_median.q +
##    slope_median.q + roughness_median.q + lake + good_snowcover +
##    nr_red_pixels.l + nr_analysed_days.l + (1 | ObsYear)
##    Data: dat
##    REML criterion at convergence: 444.1588
##    Random effects:
##    Groups Name Std.Dev.
##    ObsYear (Intercept) 22.23
##    Residual 20.75
##    Number of obs: 68, groups: ObsYear, 4
##         Fixed Effects:
##           (Intercept)         NDVI_median.l         NDVI_perc.l
##             102.6571             -21.1696             -21.7884
##         DEM_median.l         aspect_median.l         aspect_perc.l
##             72.6696               0.4308             12.3318
##         slope_median.l      roughness_median.l     NDVI_median.q
##            -180.5404              29.5651             -16.0145
##           NDVI_perc.q          DEM_median.q       aspect_median.q
##               6.8419             -60.3190             -58.8975
##         slope_median.q     roughness_median.q        lakeRitom
##              -8.0564             -52.7353              -9.0207
##              lakeTom      good_snowcoverTRUE      nr_red_pixels.l
##               2.3972              23.1105              51.2806
##    nr_analysed_days.l
##             -45.7096
```

The effect sizes of 2000 random simulations of the scrutinised variables on the first snow-free day:

##	(Intercept)	NDVI_median.l	NDVI_perc.l	DEM_median.l
## 2.5%	72.24446	-91.49612	-98.22153	-27.69055
## 50%	102.83452	-23.31642	-22.94890	75.02117
## 97.5%	132.88271	41.33982	52.19362	169.51847

##	aspect_perc.l	slope_median.l	roughness_median.l	NDVI_median.q
## 2.5%	-75.62905	-271.91999	-23.07303	-72.84198
## 50%	11.72541	-183.58041	29.31528	-16.16008
## 97.5%	92.47396	-85.89282	83.14378	39.26927

##	NDVI_perc.q	DEM_median.q	aspect_median.q	slope_median.q
## 2.5%	-50.40501	-138.441213	-125.598096	-78.373875
## 50%	6.86062	-61.451239	-60.670339	-7.828078
## 97.5%	60.71205	9.832301	3.925476	61.015388

##	roughness_median.q	lakeRitom	lakeTom	good_snowcoverTRUE
## 2.5%	-109.487567	-43.751805	-23.971503	-4.177458
## 50%	-53.213399	-8.836386	2.856653	23.435410
## 97.5%	1.744764	27.608100	28.878862	50.795889

##	nr_red_pixels.l	nr_analysed_days.l	aspect_median.l
## 2.5%	-20.76523	-188.74621	-58.504863
## 50%	51.07720	-45.23473	1.931605
## 97.5%	120.71836	88.00816	64.973136

The model output of the last snow-covered day:

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: last_snow_day_rgb   NDVI_median.l + NDVI_perc.l + DEM_median.l +
## aspect_median.l + aspect_perc.l + slope_median.l + roughness_median.l +
## NDVI_median.q + NDVI_perc.q + DEM_median.q + aspect_median.q +
## slope_median.q + roughness_median.q + lake + good_snowcover +
## nr_red_pixels.l + nr_analysed_days.l + (1 | ObsYear)
## Data: dat
## REML criterion at convergence: 382.049
## Random effects:
## Groups Name Std.Dev.
## ObsYear (Intercept) 21.255
## Residual 9.809
## Number of obs: 69, groups: ObsYear, 4
##        Fixed Effects:
##         (Intercept)        NDVI_median.l          NDVI_perc.l
##           129.1664               2.1783              -5.0095
##         DEM_median.l        aspect_median.l         aspect_perc.l
##            48.0877             -19.3215               6.0351
##        slope_median.l     roughness_median.l      NDVI_median.q
##           -36.7560              10.1839              -6.1430
##         NDVI_perc.q         DEM_median.q         aspect_median.q
##           -16.7446              16.4795               5.7870
##        slope_median.q     roughness_median.q        lakeRitom
##            -8.2110             -18.0581               0.5308
##            lakeTom      good_snowcoverTRUE       nr_red_pixels.l
##             8.8384               2.6859              -4.8940
## nr_analysed_days.l
##            31.4608
```

The effect sizes of 2000 random simulations of the scrutinised variables on the last snow-covered day:

##	(Intercept)	NDVI_median.l	NDVI_perc.l	DEM_median.l
## 2.5%	106.1597	-32.25199	-40.406346	0.1661442
## 50%	128.6404	1.25440	-5.054103	48.4727381
## 97.5%	152.0935	33.69531	29.155551	95.0270608
##	aspect_perc.l	slope_median.l	roughness_median.l	NDVI_median.q
## 2.5%	-34.070678	-80.67382	-14.28157	-33.089011
## 50%	6.603544	-36.08133	10.68173	-5.872921
## 97.5%	45.506453	7.18702	35.64512	20.026705
##	NDVI_perc.q	DEM_median.q	aspect_median.q	slope_median.q
## 2.5%	-42.408581	-20.36806	-25.194004	-42.196259
## 50%	-16.301345	16.59769	5.269092	-8.762032
## 97.5%	9.376624	50.04306	36.095215	22.920873
##	roughness_median.q	lakeRitom	lakeTom	good_snowcoverTRUE
## 2.5%	-42.533080	-16.419419	-3.226825	-10.967301
## 50%	-18.714894	0.567003	8.942644	2.712897
## 97.5%	6.929966	17.976258	20.738734	15.961362
##	nr_red_pixels.l	nr_analysed_days.l	aspect_median.l	
## 2.5%	-38.292646	-47.34355	-46.367484	
## 50%	-5.213567	30.98090	-19.065637	
## 97.5%	29.606366	109.42109	7.902201	

The model output of the number of temporary snow-free days:

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: nr_snowfree_btw_firstsnowfree_and_lastsnow_rgb   NDVI_median.l +
## NDVI_perc.l + DEM_median.l + aspect_median.l + aspect_perc.l +
## slope_median.l + roughness_median.l + NDVI_median.q + NDVI_perc.q +
## DEM_median.q + aspect_median.q + slope_median.q + roughness_median.q +
## lake + good_snowcover + nr_red_pixels.l + nr_analysed_days.l +
## (1 | ObsYear)
## Data: dat
## REML criterion at convergence: 342.7565
## Random effects:
## Groups Name Std.Dev.
## ObsYear (Intercept) 5.653
## Residual 6.943
## Number of obs: 69, groups: ObsYear, 4
##         Fixed Effects:
##          (Intercept)        NDVI_median.l        NDVI_perc.l
##               4.8929            -14.6376             0.5188
##          DEM_median.l        aspect_median.l        aspect_perc.l
##              -4.3935             -3.7027             6.7671
##          slope_median.l     roughness_median.l   NDVI_median.q
##              42.8952              5.0600            22.2699
##          NDVI_perc.q         DEM_median.q        aspect_median.q
##              -0.2039              6.2691            12.0827
##          slope_median.q     roughness_median.q      lakeRitom
##               4.1523              1.0873             8.1590
##              lakeTom   good_snowcoverTRUE      nr_red_pixels.l
##               6.2943             -3.9862            -21.1393
## nr_analysed_days.l
##              17.3877
```

The effect sizes of 2000 random simulations of the scrutinised variables on the number of temporary snow-free days:

##	(Intercept)	NDVI_median.l	NDVI_perc.l	DEM_median.l
## 2.5%	-3.785879	-38.197503	-25.1925797	-37.894418
## 50%	4.840666	-14.376137	0.3762861	-4.605041
## 97.5%	13.534692	7.553256	24.8550139	28.905782

##	aspect_perc.l	slope_median.l	roughness_median.l	NDVI_median.q
## 2.5%	-20.160954	12.14957	-12.048890	2.593939
## 50%	6.258375	43.09461	5.080051	22.119277
## 97.5%	33.893743	74.44661	22.556220	40.410760

##	NDVI_perc.q	DEM_median.q	aspect_median.q	slope_median.q
## 2.5%	-18.9398563	-17.454776	-9.279787	-19.396455
## 50%	-0.5637442	6.132915	11.992588	4.132514
## 97.5%	19.2230585	30.377628	35.115799	27.312118

##	roughness_median.q	lakeRitom	lakeTom	good_snowcoverTRUE
## 2.5%	-17.726059	-4.486809	-2.570145	-13.515116
## 50%	1.245731	8.251096	6.293327	-4.122277
## 97.5%	19.512149	20.433668	14.706896	5.638722

##	nr_red_pixels.l	nr_analysed_days.l	aspect_median.l
## 2.5%	-44.024984	-21.92965	-22.337002
## 50%	-21.280479	17.21972	-3.999674
## 97.5%	2.760643	57.23221	15.493220

6.3.6 Maximal generalized linear mixed model (GLMM) - without snow

The model output of the GLMM without snow:

```
##   Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##   Family: binomial ( logit )
##   Formula:
##   Occupied   aspect_median.z.l + aspect_median.z.q + aspect_perc.z.l +
##       DEM_median.z.l + DEM_median.z.q + NDVI_median.z.l + NDVI_median.z.q +
##       NDVI_perc.z.l + NDVI_perc.z.q + roughness_median.z.l + roughness_median.z.q +
##       slope_median.z.l + slope_median.z.q + slope_perc.z.l + z_dist_nest.z.l +
##       z_dist_road.z.l + (1 | ObsYear)
##   Data: dat
##           AIC          BIC          logLik        deviance        df.resid
##        338.1326     412.8050      -151.0663       302.1326           450
##   Random effects:
##   Groups Name Std.Dev.
##   ObsYear (Intercept) 0.00000001461
##   Number of obs: 468, groups: ObsYear, 6
##   Fixed Effects:
##           (Intercept)       aspect_median.z.l       aspect_median.z.q
##              -1.179                 -2.822                 -15.401
##        aspect_perc.z.l         DEM_median.z.l          DEM_median.z.q
##              -3.976                -23.549                 -10.650
##     NDVI_median.z.l         NDVI_median.z.q          NDVI_perc.z.l
##             -11.895                  1.903                  5.516
##        NDVI_perc.z.q    roughness_median.z.l   roughness_median.z.q
##              -8.531                  9.299                  -7.961
##     slope_median.z.l        slope_median.z.q          slope_perc.z.l
##             -14.506                -27.829                 -10.173
##     z_dist_nest.z.l         z_dist_road.z.l
##             -41.350                -29.223
```

The effect sizes of 10000 random simulations of the scrutinised variables on the occupancy:

##	(Intercept)	aspect_median.z.l	aspect_median.z.q	aspect_perc.z.l
## 2.5%	-1.6778577	-10.277090	-24.50972	-13.171639
## 50%	-1.1798365	-2.805772	-15.31037	-3.937547
## 97.5%	-0.6786148	4.534343	-6.34724	5.078711
##	DEM_median.z.l	DEM_median.z.q	NDVI_median.z.l	NDVI_median.z.q
## 2.5%	-41.904821	-28.384545	-20.640935	-5.416407
## 50%	-23.531707	-10.509052	-11.906396	1.883320
## 97.5%	-5.063823	7.472825	-3.245901	9.467227
##	NDVI_perc.z.l	NDVI_perc.z.q	roughness_median.z.l	
## 2.5%	-2.074788	-15.323435	1.595708	
## 50%	5.586149	-8.462788	9.359509	
## 97.5%	12.991886	-1.913742	17.069203	
##	roughness_median.z.q	slope_median.z.l	slope_median.z.q	
## 2.5%	-17.314035	-29.4774897	-42.03147	
## 50%	-8.036526	-14.6532249	-27.91310	
## 97.5%	1.370420	-0.1272073	-13.72066	
##	slope_perc.z.l	z_dist_nest.z.l	z_dist_road.z.l	
## 2.5%	-19.834967	-55.63442	-42.34959	
## 50%	-10.149770	-41.26129	-29.15469	
## 97.5%	-0.320693	-26.77036	-16.10073	

The computed statistical measures:

| ##	Observed data	Replicated data	Bayesian p-value
## Prop of Absence	51.5	46.58 - 56.41	0.4915
## 10% quantile of residuals	-1.02	-1.05	0.5662
## 90% quantile of residuals	0.84	0.88	0.5213

The effect plots of the strong effects detected in the GLMM without snow show the presence (index of success ~1) and absence (index of success ~0) points, the fitted model (black line), the 95% credibility interval (grey envelope) and a gridline at an index of success of 0.5.
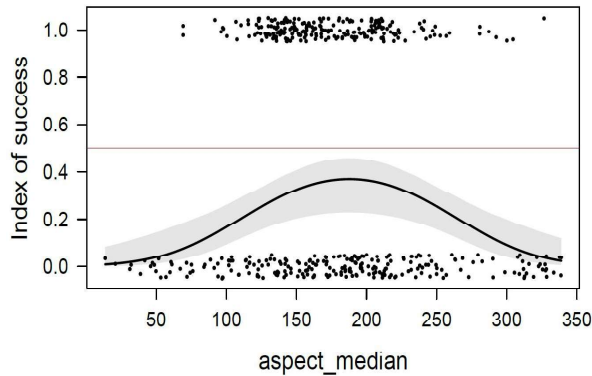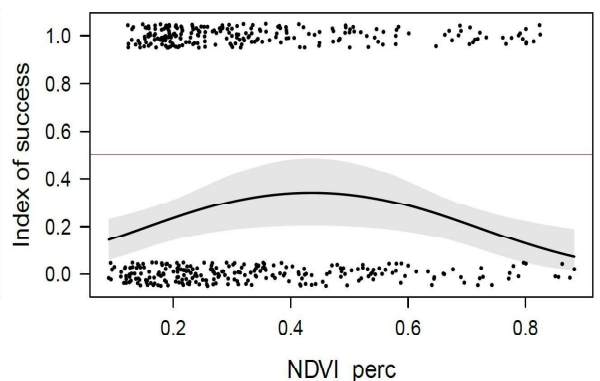

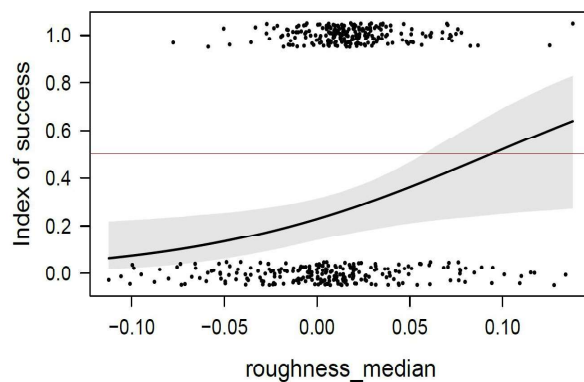(a) The median of aspect displays a maximum for S-exposed slopes.
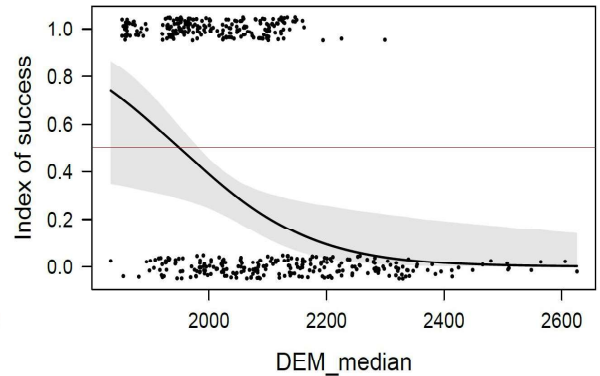
(b) The median of elevation predicted a preference for altitudes lower than 2000m a.s.l.
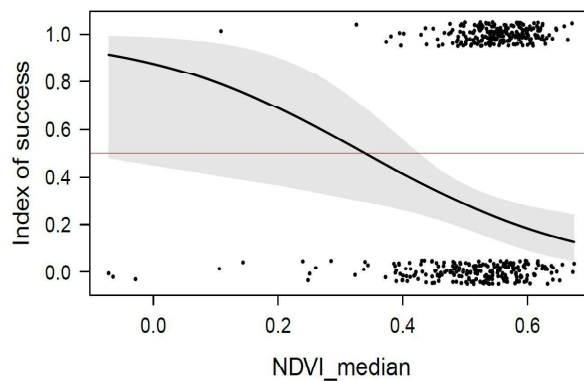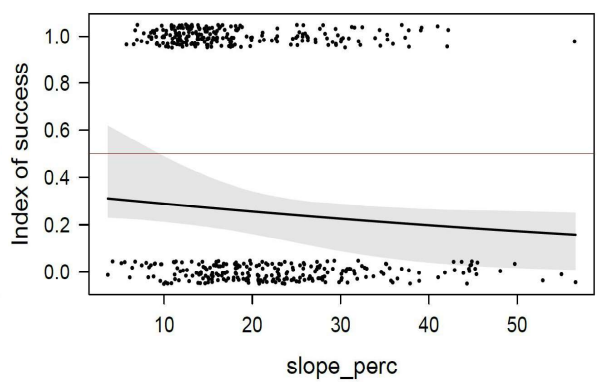

(a) The median of NDVI shows a clear preference for nest sites with no or sparse vegetation.
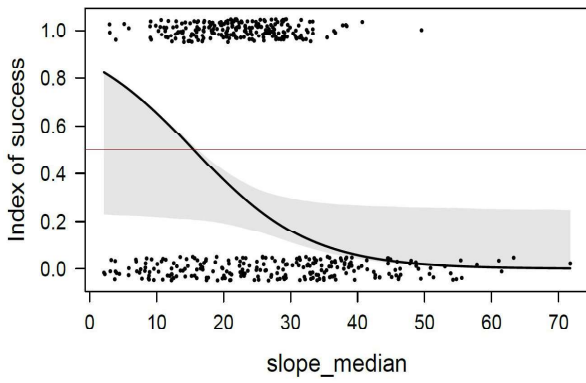

(b) The 90% percentile distance of NDVI displays a maximum for variations of 0.3 to 0.6 of the NDVI values.
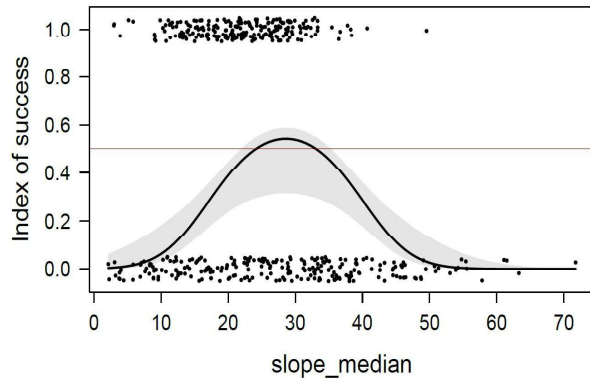

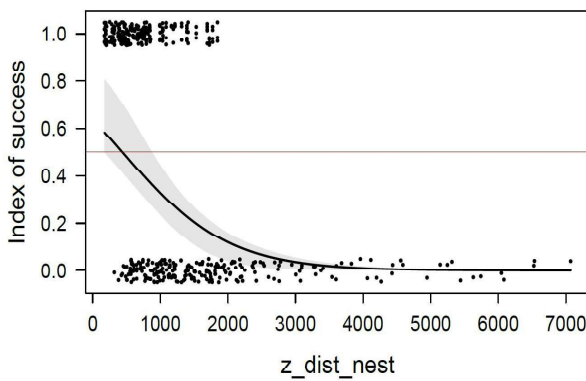(a) The median of roughness indicates a preference for values larger than 0.05.

(b) The 90% percentile distance of slope illustrates a slight preference of invariability of inclinations.
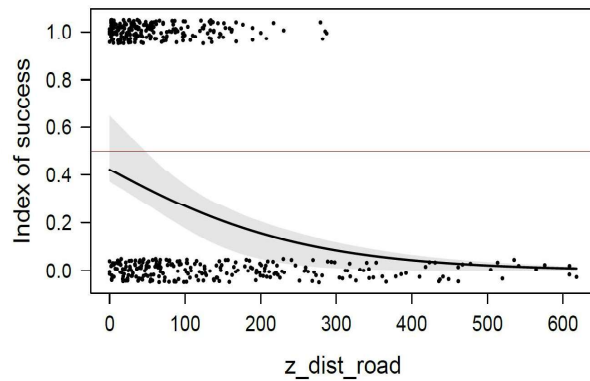
(a) The linear component of the median of slope hints at a preference of flat areas.

(b) The quadratic component of the median of slope discloses a maximum between 25° and 35°.

(a) Small cumulated distances to the three nearest nests were preferred.

(b) A proximity to the nearest path was favoured.

6.3.7 Maximal generalized linear model (GLM) - without snow

The model output of the GLM without snow:

```
##
##    Call:    glm(formula = Occupied   aspect_median.l + aspect_median.q +
##             aspect_perc.l + DEM_median.l + DEM_median.q + NDVI_median.l +
##             NDVI_median.q + NDVI_perc.l + NDVI_perc.q + roughness_median.l +
##             roughness_median.q + slope_median.l + slope_median.q + slope_perc.l +
##             z_dist_nest.l + z_dist_road.l, family = binomial, data = dat)
##
##          Coefficients:
##            (Intercept)        aspect_median.l         aspect_median.q
##                 -1.179                 -2.822                 -15.401
##           aspect_perc.l           DEM_median.l             DEM_median.q
##                 -3.976                -23.549                 -10.650
##          NDVI_median.l          NDVI_median.q              NDVI_perc.l
##                -11.895                  1.903                   5.516
##            NDVI_perc.q     roughness_median.l      roughness_median.q
##                 -8.531                  9.299                  -7.961
##         slope_median.l         slope_median.q             slope_perc.l
##                -14.506                -27.829                 -10.173
##          z_dist_nest.l          z_dist_road.l
##                -41.350                -29.223
##
##    Degrees of Freedom: 467 Total (i.e. Null); 451 Residual
##    Null Deviance: 648.4
##    Residual Deviance: 302.1 AIC: 336.1
```

The effect sizes of 10000 random simulations of the scrutinised variables on the occupancy:

##	(Intercept)	aspect_median.l	aspect_median.q	aspect_perc.l
## 2.5%	-1.6801725	-10.108278	-24.19861	-13.524320
## 50%	-1.1768810	-2.766928	-15.34347	-3.990701
## 97.5%	-0.6695571	4.403221	-6.59559	5.265106

##	DEM_median.l	DEM_median.q	NDVI_median.l	NDVI_median.q
## 2.5%	-42.185655	-28.851511	-20.643137	-5.600604
## 50%	-23.342355	-10.478405	-11.808806	1.804540
## 97.5%	-4.762709	7.008469	-3.093634	9.389051

##	NDVI_perc.q	roughness_median.l	roughness_median.q	slope_median.l
## 2.5%	-15.166356	1.730467	-17.371028	-29.3182239
## 50%	-8.431196	9.317223	-7.890525	-14.5357249
## 97.5%	-1.692467	16.831585	1.715466	-0.1341736

##	slope_median.q	slope_perc.l	z_dist_nest.l	z_dist_road.l
## 2.5%	-42.18328	-19.9381476	-55.50878	-42.35401
## 50%	-27.88002	-10.1757503	-41.28602	-29.32922
## 97.5%	-13.43523	-0.2265823	-26.91647	-16.26163

##	NDVI_perc.l
## 2.5%	-2.264284
## 50%	5.501760
## 97.5%	13.120322

The computed statistical measures:

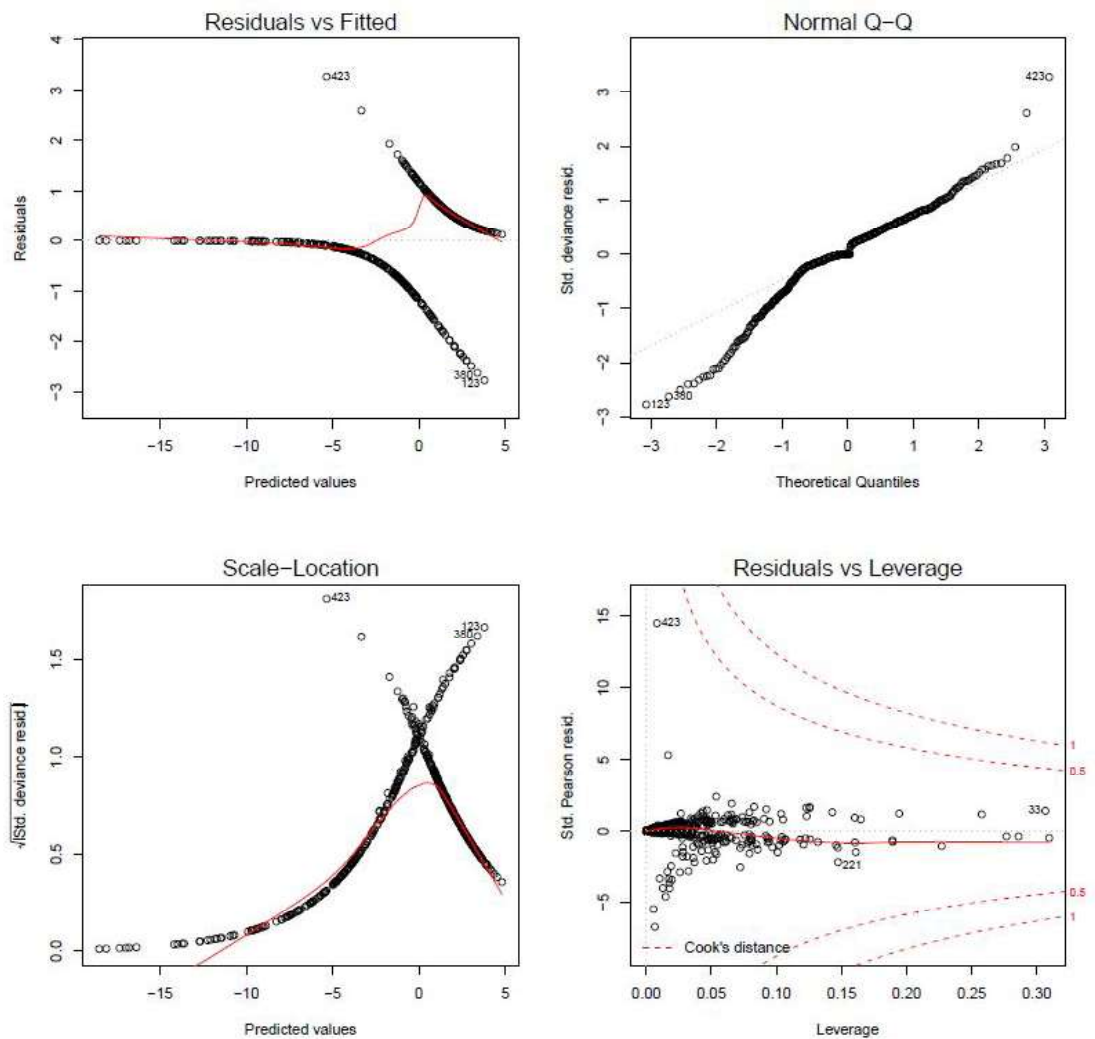| ##	Observed data	Replicated data	Bayesian p-value
## Prop of Absence	51.5	46.58 - 56.2	0.4886
## 10% quantile of residuals	-1.02	-1	0.5588
## 90% quantile of residuals	0.84	0.85	0.5136

Figure 82: The created residual plots in the top row are the Tukey-Anscombe plot (left) and the normal Quantile-Quantile plot (Q-Q plot) of residuals, in the bottom row are the scale-location plot (left) and the leverage plot (right) [F. Korner-Nievergelt, Roth, et al. 2015]. The mean of the residuals is not always around zero in the Tukey-Anscombe plot, which is acceptable for a binomial distribution [F. Korner-Nievergelt and P. Korner-Nievergelt 2019]. The normal distributed residuals lay on the diagonal in the normal Q-Q plot [F. Korner-Nievergelt, Roth, et al. 2015]. The analysis of the normal Q-Q plot of a binomial model is shown in Figure 84. The smoother in the third panel is horizontal if the residuals display a homogeneous variance [F. Korner-Nievergelt, Roth, et al. 2015]. The here displayed variance is not homogeneous yet acceptable for a binomial model. The last panel displays a higher leverage for observations with a rare combinations of factors and shows different leverage values in for an unbalanced dataset [F. Korner-Nievergelt, Roth, et al. 2015]. Furthermore, the cook's distance measures the influence of an observation on the data and is >1 for influential observations [F. Korner-Nievergelt, Roth, et al. 2015]. The here displayed data seems balanced and without influential observations.
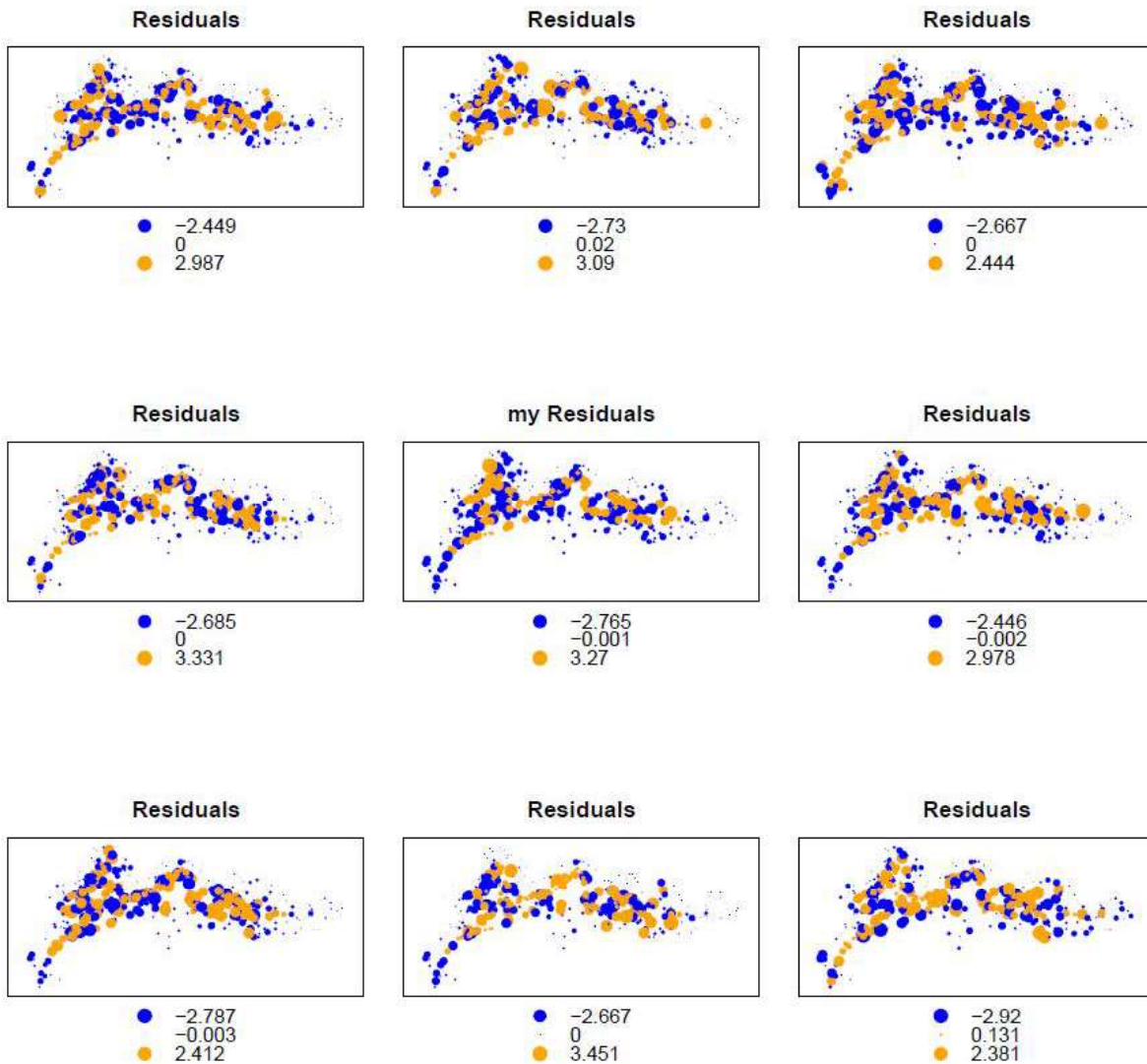
Figure 83: The residuals of my fitted GLM are shown in the central plot and display a spatial autocorrelation, i.e. a pattern in the spatial distribution of residuals is detectable [F. Korner-Nievergelt and P. Korner-Nievergelt 2019]. The surrounding plots origin in simulated plots. The spatial distribution of my model is not distinguishable from the simulated distributions. Therefore, the present spatial autocorrelation is explained and needs no further precautions [F. Korner-Nievergelt, Roth, et al. 2015].
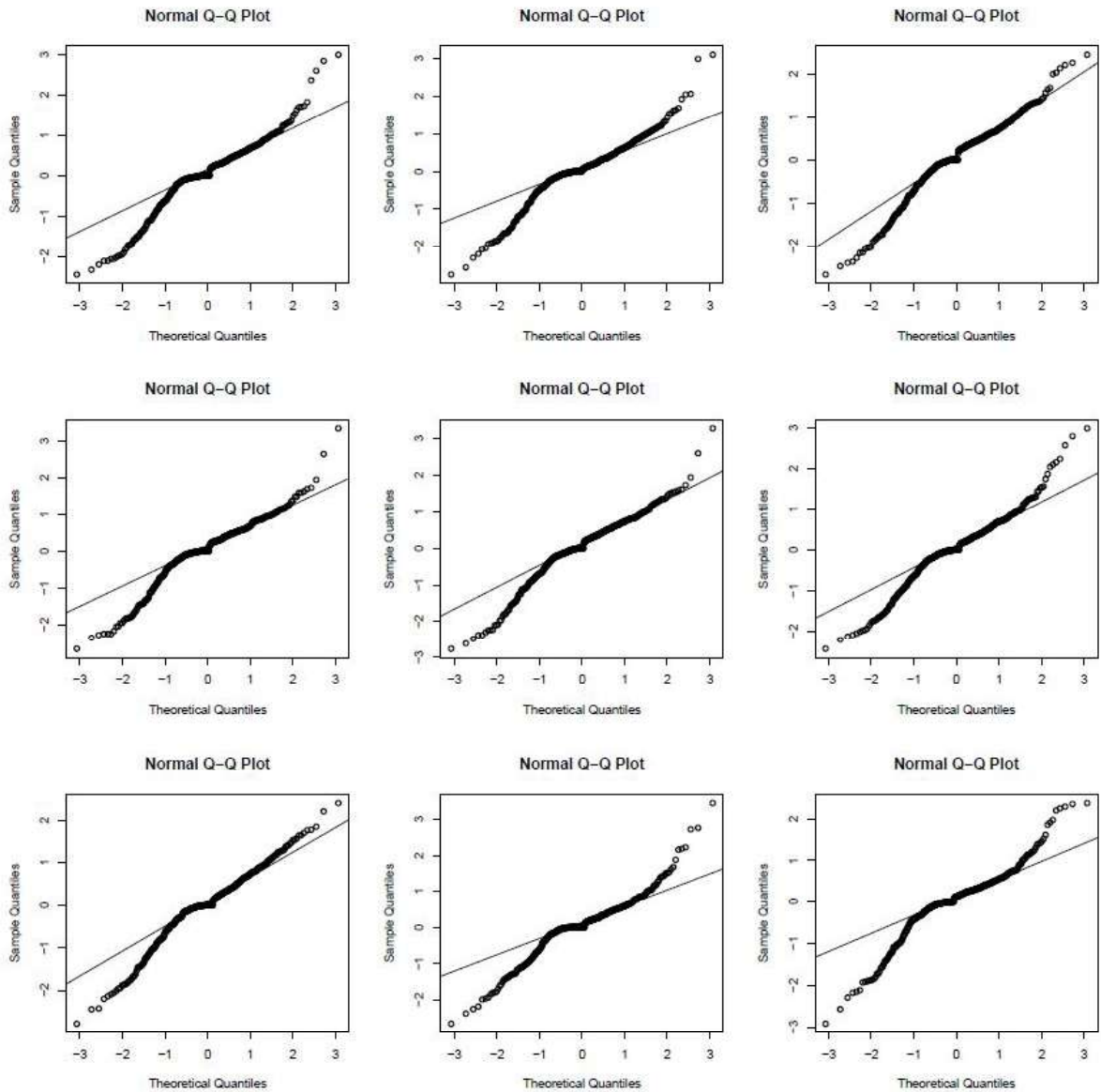
Figure 84: The central Q-Q plot origins in my fitted GLM while the surrounding Q-Q plots display simulated values. If the Q-Q plot of my model can be distinguished from the simulated plots, its residual deviation deviates from normal [F. Korner-Nievergelt, Roth, et al. 2015]. The central Q-Q plot here cannot be distinguished visually from the surrounding ones. Therefore, the assumption of normality is not violated [F. Korner-Nievergelt, Roth, et al. 2015].

6.3.8 Maximal generalized linear model (GLM) - with snow

The model output of the GLM with snow:

```
##
## Call:    glm(formula = Occupied   aspect_median.l + DEM_median.l + DEM_median.q +
##              NDVI_median.l + slope_median.l + slope_median.q + roughness_median.l +
##              z_dist_nest.l + z_last_snow.l + z_first_snowfree.l + DEM_median.l:slope_median.l,
##              family = binomial, data = dat)
##
##         Coefficients:
##             (Intercept)                   aspect_median.l
##                  -3.559                             9.543
##          DEM_median.l                      DEM_median.q
##                 -28.577                            -5.591
##          NDVI_median.l                     slope_median.q
##                   6.246                           -52.160
##          slope_median.q                  roughness_median.l
##                 -51.971                            -3.204
##          z_dist_nest.l                      z_last_snow.l
##                 -12.712                             2.963
##      z_first_snowfree.l    DEM_median.l:slope_median.l
##                  -1.747                           369.734
##
##      Degrees of Freedom: 67 Total (i.e. Null); 56 Residual
##      Null Deviance: 93.74
##      Residual Deviance: 24.78 AIC: 48.78
```

The effect sizes of 10000 random simulations of the scrutinised variables on the occupancy:

##	(Intercept)	aspect_median.l	DEM_median.l	DEM_median.l: slope_median.l
## 2.5%	-7.3696862	-3.900603	-59.931622	-5.226838
## 50%	-3.5475887	9.603872	-28.638032	369.270727
## 97.5%	0.2438729	22.736556	2.294206	754.568244
##	DEM_median.q	NDVI_median.l	roughness_median.l	slope_median.l
## 2.5%	-27.008664	-18.320690	-16.072101	-89.95644
## 50%	-5.562695	6.192385	-3.177537	-52.25082
## 97.5%	16.166461	30.841214	9.614017	-14.01641
##	slope_median.q	z_dist_nest.l	z_first_snowfree.l	z_last_snow.l
## 2.5%	-91.24264	-27.283674	-15.621428	-13.988515
## 50%	-51.79708	-12.669151	-1.777085	3.074934
## 97.5%	-12.57068	1.787414	12.277490	20.136326

The computed statistical measures:

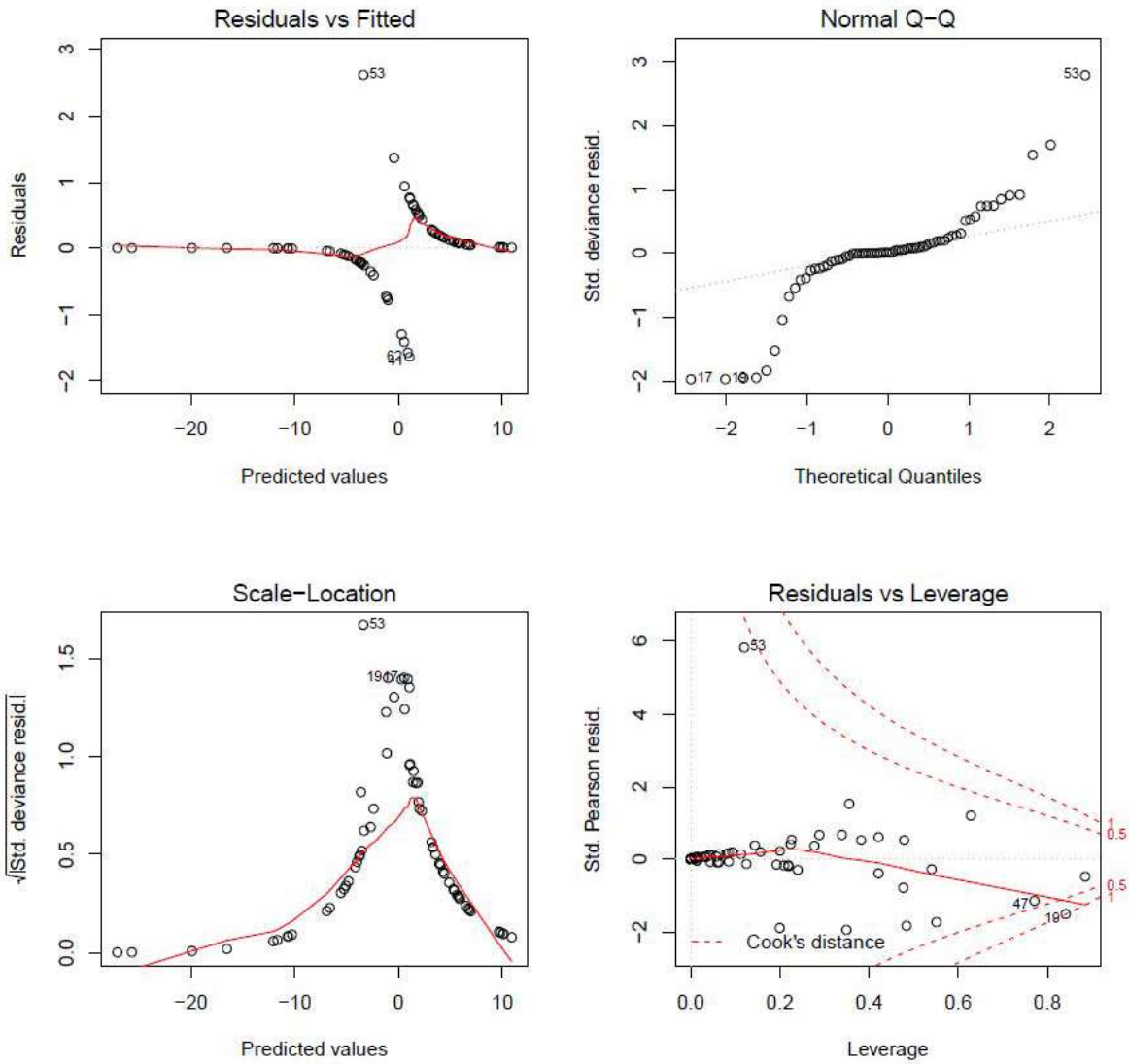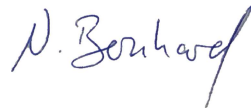| ##	Observed data	Replicated data	Bayesian p-value
## Prop of Absence	45.59	32.35 - 57.35	0.4936
## 10% quantile of residuals	-0.51	-0.03	0.3229
## 90% quantile of residuals	0.61	0.15	0.4492

Figure 85: The same residual plots as for the GLM without snow were computed. The Tukey-Anscombe plot displays an acceptable distribution despite the extreme observation 41, 53 and 62. The smoother in the third panel deviates from a homogenous variance in residuals. The cook's distances never exceed one, which reveals no violations of the model assumptions.

7 STATEMENT OF AUTHORSHIP

Personal declaration: I hereby declare that the submitted thesis is the result of my own, independent work. All external sources are explicitly acknowledged in the thesis.

Winterthur, 30/06/2019   Nadja Bernhard