Carl von Ossietzky Universität Oldenburg
Fakultät V
Mathematik- und Naturwissenschaften
IBU Institut für Biologie und Umweltwissenschaften
Studiengang Diplom-Landschaftsökologie
AG Landschaftsökologie

DIPLOMA THESIS

Habitat selection of the Sociable Lapwing *Vanellus gregarius* in Central Kazakhstan – a modelling approach

submitted by:	Johannes Kamp
subject:	Landscape Ecology (Dipl.)
supervisor:	PD Dr. Robert Biedermann, Oldenburg
second supervisor:	Dr. Robert Sheldon, London

Oldenburg, October 2007

This diploma thesis is part of the International Sociable Lapwing Research and Conservation project by BirdLife International, The Royal Society for the Protection of Birds (RSPB) and The Association for the Conservation of Biodiversity in Kazakhstan (ASBK). Fieldwork was funded by the German Academic Exchange Service (DAAD), Defra/The Darwin Initiative, the RSPB and Foerderkreis fuer allgemeine Naturkunde/Biologie (Jena/Sheffield).

CONTENTS

1 Introduction

2 Materials and Methods

3 Results

4 Discussion

5 Conclusions and perspectives

7 Abstracts

APPENDICES

List of figures

List of tables

„Am 27. April hatten wir die Freude, die ersten Herdenkiebitze (*Vanellus gregarius*) zu sehen, und zwar hielten sie sich, ihrem Namen treu bleibend, bei und unter den Viehherden auf, wie dies mit Vorliebe ihre Gewohnheit ist.“

(from: Finsch, O. 1879: Reise nach West-Sibirien im Jahre 1876. Wallroth, Berlin.)

1 INTRODUCTION

1.1 Background and objectives

The breeding population of the Sociable Lapwing *Vanellus gregarius* [=*Chettusia gregaria*] is endemic to the vast steppes of Kazakhstan and Southern Russia. Due to a large scale historic and accelerated recent decline in population numbers, the species' threat status was recently updated to "Critically Endangered" in the world Red Data Book (BIRDLIFE INTERNATIONAL 2006). Between 2000 and 2002, alarming news on decreasing numbers from various study areas across the distribution range were published. The size of the world population was estimated as 5,000 pairs at the end of the 1980s (COLLAR et al. 1994), but already in the year 2000, KHROKOV (2000) suggested overall population size to hardly exceed 1,000 pairs. This alarming development led to the organization of an expert workshop on Sociable Lapwing and other threatened steppe waders in Moscow 2002. During the workshop, a detailed Species Action Plan was compiled (TOMKOVICH and LEBEDEVA 2004), comprising suggestions for better conservation of the species, which should be implemented both internationally and on a national level. The action plan identified an urgent need to conduct sound research at the breeding grounds aiming to collect data on the reasons of the decline and suggesting conservation measures. A new world population estimate of only 171-516 remaining breeding pairs was presented.

As a first step in Action Plan implementation, a research project on the breeding grounds of Sociable Lapwing was initiated in the Korgalzhyn region, Central Kazakhstan, in 2004. This project was launched by BirdLife International and has been realized by The Royal Society for the Protection of Birds (RSPB) and the Association for the Conservation of Biodiversity in Kazakhstan (ACBK). By the means of nest and chick survival monitoring, radio tracking and colour ringing, factors influencing breeding success and population development in the breeding areas should be determined and measures for habitat management should be developed (WATSON et al. 2006, SHELDON et al. 2005). This diploma thesis was prepared in the framework of the above mentioned project, focusing on research on habitat and nest site selection (as defined by JONES 2001) of the Sociable Lapwing. Sound quantitative knowledge on this topic is crucial for the implementation of any future habitat management measures.

The main aims and objectives of the study are:

(i) to define the main factors influencing habitat and nest site selection of the Sociable Lapwing,
(ii) to provide a data base for predictive distribution models for more effective survey work, and
(iii) to develop practical measures for habitat management based on the new findings.

1.2 Basic species' ecology info

The Sociable Lapwing is a steppe and semi-desert species concentrated in areas, where short vegetation is available, endemic to the western Eurasian steppe zone of South Russia and Kazakhstan. There are no breeding records from the Mongolian steppes. The species breeds in small, loose colonies of 5(3) to 15 (30) pairs, preferably in the vicinity of grazing livestock (FINSCH 1879, DOLGUSHIN 1962). Areas inhabited by Sociable Lapwing are usually species-poor, but rather often, White-winged and Black Lark *Melanocorypha leucopterus* and *yeltoniensis* breed at the same sites (pers. obs., Appendix D-1, D-3, D-7). The species seems to be ecologically close to Northern Lapwing *Vanellus vanellus* and is often socialized with it, both in the breeding area and during migration, although it appears to be less dependent on wetland habitats. Taxonomic status is still unclear, but placing it into the *Vanellus* group (was earlier treated as *Chettusia gregaria*) is probably more consistent than grouping it with the smaller plovers.

Maturity may be delayed to the second calendar year in some cases, but birds are capable to breed in their first year, as recently proved by colour ringing methods (SHELDON et al. 2005). Display and mating takes place already during the spring migration and many birds arrive paired, but others pair up after the arrival in the breeding areas. Colonies are established and nest sites selected during 10-15 days after the birds' arrival (DOLGUSHIN 1962). Clutches usually consist of four eggs, but birds are capable of producing five eggs (pers. obs., Appendix D-8, D-9). RYABOV (1949) recorded four eggs in 81 % and five eggs in 19% in a sample of 66 nests at Naurzum, North Kazakhstan. For detailed data on breeding ecology and productivity, cf. WATSON et al. (2006) and SHELDON et al. (2005). Both RYABOV (1949) and results from the mentioned papers point to a mean incubation period of 25 days, ranging from 24 to 26. Most males leave the females after clutches hatch and gather in flocks, but some individuals stay until chicks have fledged (pers. obs.).

The species is mainly active during daytime, but there is little information available on nocturnal behaviour. Food is collected by probing and pecking at soil surface and vegetation. Main food items include beetles, locusts and spiders. Analysis of stomach contents in Naurzum, North

Kazakhstan, revealed high amounts of Curculionidae and Tenebrionidae in May, *Acrida*, *Omocestus* and *Locusta* grasshoppers from mid June to August, and beetles of the families Carabidae, Elateriade, Chrysomelidae, Curculionidae and Tenebrionidae in September (RYABOV 1949). The amount of vegetable food is negligible.

Phenology: Birds arrive at the breeding areas from mid April to the beginning of May. Chicks hatch mid May to mid July (Appendix D-10), and fledge mid July to end August. Post breeding flocks (with up to 150, sometimes 1,000 birds) gather in July and August. The birds leave the breeding grounds from the beginning of August onwards in Northern Kazakhstan, with the last individuals observed in early September. The wintering areas are probably reached in November-December, birds are then vagrant throughout larger areas. Departure for the breeding grounds starts already in the beginning of February (for timetable details cf. DOLGUSHIN 1962, TOMKOVICH and LEBEDEVA 2004).

1.3 Distribution and migration routes

The Sociable Lapwing was once a widespread breeder across the Eurasian steppe belt (Fig. 1). The westernmost breeding grounds were situated in Ukraine and SW Russia. The species disappeared from these areas before the end of the 19th century due to agricultural development with steppe ploughing, but single pairs were observed in Ukraine until the 1960s (SHARLEMAN 1938, DOLGUSHIN 1962) (Fig. 1). The easternmost known breeding occurrences are from Xinjiang Autonomous Region, China (FINSCH 1879, GENSICHEN 1982). There are no Chinese breeding records after the 1870s (M. MING in litt.) and no records of migrating birds after 1980.

A current data base project reviewing worldwide distribution and migration routes (KOSHKIN and KAMP in prep.), supported by the results of rapid surveys over large territories in 2005 to 2007 (SHELDON et al. 2006), came to the following conclusions:
The current breeding distribution of Sociable Lapwing stretches between Eastern Kazakhstan (Saissan lake area) and SW Russia (ca. 49°50’ E to 83°30’ E). Highest densities are reached within the terrestrial ecozone “Kazakh steppe” (OLSON et al. 2001). I estimate that currently 80% of the breeding population are concentrated within the latter region. It is unlikely that there are current breeding colonies west of the Volga river. The northern distribution border follows the northern forest steppe belt border (OLSON et al. 2001, Fig. 1). The southern border of breeding distribution is poorly studied. Birds breed in semi-deserts north of the Caspian Sea, the Aral Sea, and in Central Kazakhstan southward to at least 48°30’ N. A recent breeding record from the Taukum desert in Southern Kazakhstan at 44°30’ N, 75°50’ E (BELYALOV 2004) suggests that breeding colonies south of the steppe and semi-desert zone in Northern and Central Kazakhstan might have been largely overlooked.

Migration probably takes place on “narrow broad front”, i.e. there are no narrow corridors, but distinct migration routes probably linked with the availability of stopover sites. According to the review of nearly 1,000 data sets, ring recoveries and results of satellite tracking (KOSHKIN and KAMP in prep.), three of such are proposed. The majority of individuals heads west at the end of July and beginning of August, probably flying parallel to degrees of latitude. These birds arrive at the steppe and semi-desert areas of the Northern Caucasus in late August and stay throughout September. At these autumn stopover sites, concentrations of up to 1,000 resting birds have been observed both 150 years ago and recently (RADDE 1884, MALOVICHKO et al. 2006). There, they have good opportunities to refuel due to a range of high-quality wetlands and possibly also through utilising arable fields. Birds then depart when fit for a non-stop, high-altitude flight crossing the Caucasus range and Eastern Turkey with the Middle East to north eastern Africa. These birds originate from areas situated as easterly as 70°00’ (proved by colour ring recoveries and satellite tracking, SHELDON et al. 2006, SHELDON pers. comm. 2007). Passage is inconspicuous in most places, but regularly some larger flocks rest at the steppes of eastern Turkey and northern Syria in February/March and October. Recent records concern e.g. 1,500 birds in Syria, February 2007, and more than 1,000 birds at Ceylanpinar IBA, Turkey in October 2007, R. HOFLAND, Ö. BALKIZ pers. comm., cf. also BIRDLIFE INTERNATIONAL 2007). These accumulations might indicate a last spring migration stopover site before more or less non-stop flight to the breeding grounds, although local shepherds stated that larger flocks winter at the sites (BIRDLIFE INTERNATIONAL 2007).

The Sociable Lapwings using this route winter partly in Israel, but counts there account only for a small proportion of the suggested number of birds using this flyway. Main wintering areas were situated in Sudan and Eritrea prior to 1950, but the current distribution is largely unknown. There is evidence that most birds still winter in the Sahelian Acacia Savannah (OLSEN et al. 2001) of Sudan, Northern Eritrea and maybe Chad. Historic incidence of birds wintering in Somalia and Ethiopia can currently not be supported. Conspicuous passage during both spring and autumn in Azerbaidzhan (flocks of up to 180 birds in 2007, K. GAUGER, J. ETZOLD, J. PEPER pers. comm.) probably refers to birds of this flyway.

The “central flyway” comprises birds from Central as well as Eastern Kazakhstan. These depart via Southern Kazakhstan, and continue via Uzbekistan, Western Afghanistan and Iran to wintering areas along the Arabian Gulf coast in Southern Iran, the United Arab Emirates and Oman. In winter 2006/07 there were approximately 30 birds present synchronously at known sites in the mentioned countries (A. KHALEGHIZADEH, H. VAN DIEK, T. PEDERSEN and H. KUNZE pers. comm.)

Birds moving on the “eastern flyway” probably originate from the easternmost parts of the distribution range, pass through Southern Kazakhstan, Kyrgyzstan, Tadzhikistan and Eastern Afghanistan and winter at wetlands in NW India and S Pakistan. In winter 2006/07 there were approximately 50 birds present synchronously at known sites in India and Pakistan (N. DEVASAR, M. MALIK, G.R. MUGHAL pers. comm.)

Judging from the observations of the latest five years, I estimate 85% of the total population to currently use the western flyway, possibly not more than 5% migrate on central routes and the remaining 10 % use the eastern flyway to Pakistan and India.

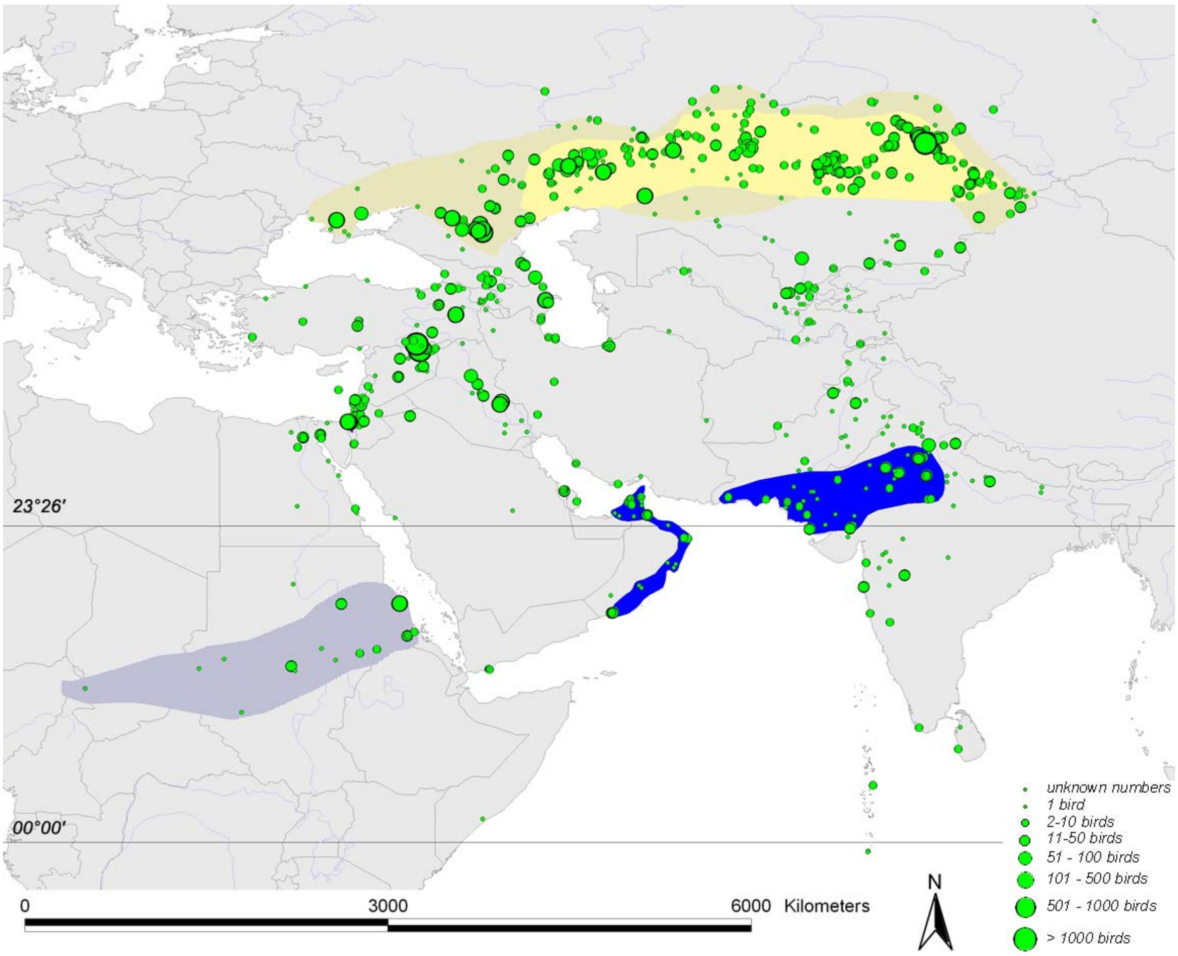


Fig. 1: Current (yellow) and former (hatched yellow) breeding distribution, known current (blue) and former/putative wintering areas (hatched blue) of Sociable Lapwing. Green dots represent sites where Sociable Lapwings have been observed between 1850 and present (n=1,017 records). Own analysis from records of the Sociable Lapwing record data base.

1.4 Causes of decline

The Sociable Lapwing population underwent serious declines over the last centuries (cf. 1.1). This becomes most obvious when comparing flock size of post breeding and migrating flocks, which has reduced from 8,000-10,000 observed in Kazakhstan still regularly around 1900 (Dolgushin 1939) to currently rarely more than 300 birds in a single post-breeding flock. The likely main cause for this until 1960 was habitat destruction and fragmentation due to steppe cultivation by the means of ploughing. This process was underway in the 18th century in Ukraine and European Russia, and in the 19th century in Asian Russia and Kazakhstan. Probably all declines and shifts of distribution borders before 1960 can be explained by these measures (Dement’ev and Gladkov 1951, DOLGUSHIN 1962, RYABOV 1974). After 1960, reasons for the strong decline have not been clarified yet. At some sites, a significant decline in both breeding and roosting numbers in the 1970s and 1980s was followed by further decreases in numbers after 1991 (e.g. RYABOV 1974, BEREZOVIKOV et al. 1998, EICHHORN and HEINECKE 2000, EICHHORN and KHROKOV 2002).

Decline in ungulate numbers and habitat deterioration A sharp decline in state livestock numbers was observed after the crash of the Soviet Union in 1991 both in Russia and Kazakhstan due to large scaled emigration of the rural human population to the cities and the collapse of the subsidized state farm system. Sheep and cattle numbers decreased by more than 80% in some regions in Kazakhstan and by 50% on average (GOSKOMSTAT 2001, ROBINSON and MILNER-GULLAND 2003, LENK 2001, WRIGHT 2002). This resulted in a strong reduction of short grazed areas, as former cattle and sheep pastures were abandoned. Changes in vegetation height and structure made many areas unsuitable as Sociable Lapwing breeding habitat (photos comparing situation in 1989 and 2005 in MOROZOV 2005). It is likely, but not yet proved, that abundance of wild ungulates influences features of the Sociable Lapwing's habitats (WATSON et al. 2006, SCHIELZETH 2005). The severe decline of the Saiga antelope *Saiga tatarica* in the past decades may therefore be linked to decreases in Sociable Lapwing numbers. Saiga numbers dropped from over a million animals in 1985 to a few more than 30,000 in 2004 due to severe poaching (BEKENOV et al. 1998, MILNER-GULLAND et al. 2001, Yu. GRACHEV in litt.). It seems possible that the bird was affected by this decline in remote areas, where no surveys have been carried out. This may have led to a complete abandonment of some areas where domestic grazing did not compensate for the decline in wild ungulate grazing.

Low breeding success Several authors suggest that eggs were destroyed frequently by livestock, dogs and humans at places close to settlements (WATSON et al. 2006, SOLOMATIN 1997). The situation got probably worse after 1991, when the numbers of private livestock raised due to a worse general food supply of the human population and livestock was stronger concentrated around the villages, but were already a problem from the beginning of the development period in Kazakhstan in the 1950s, when new cattle breeding farms were erected at several state farms (RYABOV 1974, SOLOMATIN 1997).

The increased numbers of Corvids, especially Rooks *Corvus frugilegus* due to improved nesting opportunities in maturing shelter-belts have been made responsible for decreased breeding success (BELIK 2005), but there is currently little evidence to substantiate this (SHELDON et al. 2005, R. SHELDON pers comm. 2006). In Central Europe, Corvids have a fairly low impact on breeding success of Lapwings and other open country waders, a more serious problem are carnivore mammals like foxes *Vulpes vulpes* and mustelids *Mustela erminea, putorius* (LANGGEMACH and BELLEBAUM 2005, TEUNISSEN et al. 2005).

Agricultural intensification in the breeding areas After several failures of the Virgin Land campaign, farming was increasingly mechanized and intensified in the Soviet Union (WEIN 1983). Side-effects on steppe land seem to be possible. Breeding attempts on agricultural fields suffer from high nest loss and increased chick mortality rates due to the abundant and mechanized land working (MOROZOV 2005).

Use of anti-locust chemicals It is possible, that the use of chemicals influenced feeding and foraging success of the Sociable Lapwing, but there is little data to support this. During the first years after the political changes in 1991, locust control was limited due to a lack of resources. However, by the year 2000 anti locust actions involved spraying Dimilin pesticide on 8.1 million hectares of crops, fallows, pastures and abandoned crop land (IPP CONSULTANTS 2003). Dimilin is a broad-spectrum insecticide inhibiting cuticula growth in insects and therefore having a skinning-thwarting, lethal impact. Serious side-effects on other arthropods than the targeted are denied in several papers, but may be possible. Studies on avifauna changes in forests, where Dimilin was used to control Gypsy Moths *Lymantria dispar*, revealed no treatment effects on bird abundance (COOPER et al. 1990), but bird diets (COOPER et al. 1990, SAMPLE et al. 1993), foraging behaviour (COOPER et al. 1990), and body condition (WHITMORE et al. 1993) were all affected. It is unclear, if these results may be transferred to steppe bird communities.

CHILDEBAEV (2003) noticed a significant to dramatic decrease in arthropod numbers after experimental spraying on sample plots within fallow and steppe areas using different pesticides at study sites in Northern Kazakhstan (Akmolinskaya oblast'). He detected a decline in numbers of several bird species as well, most significantly Skylarks *Alauda arvensis* were concerned. He does not lead back the shift in numbers to a higher mortality of the birds, but to a diminished food availability and large-scaled bird flight caused by that. No impact on breeding success has been traced in this investigation, the main reason for that is considered to be the late date of spraying (end of June), when most young steppe birds are nearly fledged. There are hints, that spraying during the early hatching period may cause bigger loss of juveniles (e.g. MULLIE and KEITH 1993 for African savannahs).

Climate change There is little scientific data on the effects of climatic factors on Sociable Lapwing distribution or population decline, although such links cannot be discounted (SHEVCHENKO 1998 for Volga-Ural interfluve). It seems possible that an increased frequency of droughts has affected the breeding success of the Sociable Lapwing in a negative way in some semi-desert populations. On the other hand, a stronger atlantic touch of the climate in recent years (hints for Karagandy region, Kazakhstan, ROBINSON and MILNER-GULLAND 2003) with frequent downpours might have affected nest and chick survival in recent years (BELIK 2005). An increase in the frequency of moist years might have influenced vegetation growth and accelerated the “degeneration” of former grazed areas.

Habitat loss and poaching in the wintering areas Very little information is available on this topic. More frequent droughts and wetland destruction could be responsible for a decrease in Sociable Lapwing numbers (KANNAN 2004). At least in India, the southern border of regular winter observations shifted from app. 20°N to 28°N. In Africa, increasing poaching due to increased poverty of the rural population after armed conflicts in Eritrea, Sudan and Chad could possibly influence Sociable Lapwing survival. Recently, poachers in Syria hunted on large congregations of Sociable Lapwing (BirdLife International 2007). Consistent with this hypothesis, return rates of birds ringed in Central Kazakhstan are fairly low (Sheldon et al. in prep.).

1.5 Overview of information available on habitat use

To date, there has been little quantitative research on habitat selection by Sociable Lapwing. All reports and papers available on habitat use do not quantify the influence of single parameters and are largely based on anecdote. The recent publication of Watson et al. (2006) is of preliminary character, and number of factors considered as well as sample size was small.

The Sociable Lapwing seems to be adapted to a variety of different short-vegetated habitat types in steppe and semi-desert zone. However, during the last decades colonies were found mostly on short, intensively grazed (sometimes overgrazed) pastures in the direct vicinity to settlements both in Kazakhstan and Russia. The following overview of knowledge on habitat use was compiled from the key references (Table 1).

Table 1. Information on Sociable lapwing habitat use compiled from the key references

habitat type	habitat characteristics	soil features	Site	key references
short-grass steppe	primary/"pristine" feather-grass (*Stipa* spp.) steppe, often strongly grazed	sand, often Solonets, Kastanozems	Akkumy Sands, Western Kazakhstan oblast'; Korgalzhyn region, Akmolinskaya oblast'	Belik 2005, Khrokov 1977, Sushkin 1908, Ryabov 1949, Dolgushin 1962
	strongly grazed *Artemisia-/Festuca*-steppe close to settlements	„clay“, mostly Solonets	Aksai village, Western Kazakhstan oblast'; Naurzum region, Kostanaiskaya oblast'; Korgalzhyn region, Akmolinskaya oblast'	Berezovikov et al. 1998, Dolgushin 1962, Gordienko 1991, Khrokov 1977, 1996, Ryabov 1974, Eichhorn and Heinicke 2000, Eichhorn and Khrokov 2002, Watson et al. 2006, Solomatin 1997
	at solonchaks, close to wet patches, in areas covered by haline vegetation such as salt tolerating grasses, *Camphorosma* and *Salicornia*	„clay“	Naurzum region, Kostanaiskaya oblast'; Korgalzhyn region, Akmolinskaya oblast'	Khrokov 1977, Gordienko 1991
semi-desert	sparse A*rtemisia* coverage, covered with pebbles, "very stony"; sometimes at "artesian wells" (wetter depressions) in semi-desert	„clay“	Volga-Ural-interfluve, Western Kazakhstan oblast'; Zaissan depression, Eastern Kazakhstan oblast'; Chelyabinskaya oblast', Russia	Berezovikov et al. 1998, Karyakin and Koslov 1999, Shevchenko 1998, Shevchenko et al. 1993
agriculturally used fields	recently ploughed, or covered with sparse, short corn vegetation	Kastanozems	Chelyabinskaya oblast', Russia; Orenburzhskaya oblast', Russia	Karyakin and Koslov 1999, Morozov 2005, Ili'chev and Fomin 1979, Solomatin 1997
fallow fields	dominant plant genera *Artemisia, Festuca sulcata, Salicornia, Atriplex, Sonchus, Euphorbia, Achillea, Convolvulus*		Chelyabinskaya oblast', Russia; Naurzum region, Kostanaiskaya oblast'; Korgalzhyn region, Akmolinskaya oblast'	Karyakin and Koslov 1999, Khrokov 1996, Khrokov and Karpov 1998, Solomatin 1997

2 MATERIALS AND METHODS

2.1 Study area

Fieldwork was carried out between 23rd April and 31st July 2006 in Central Kazakhstan. The study area is situated at the Lake Tengiz depression some 120 km SW of Astana, the Kazakh capital (Fig. 2). The village of Korgalzhyn constitutes the larges settlement. The study area covers an area of approximately 200 km x 150 km (30,000 km²) and stretches between 49°40' – 50°55' N and 68°38' – 70°59' E. For a more detailed geographical description of the study area cf. SCHIELZETH et al. (2007, in press).

Fig. 2. Location of the Sociable Lapwing project study area in Central Kazakhstan

The climate within the region is extremely continental: summers are hot, winters are long and cold, and precipitation peaks in summer (July/August). Spring often lasts not longer than two weeks, September and October are typical autumn months with greater cloud coverage and cool temperatures. At Korgalzhyn, mean annual temperature is 2.8°C, mean July temperature +20° and mean January temperature -19°C. Maximum temperature amplitude is 94 K (-51° to +43°C), mean temperature amplitude 72 K (-32 to +40°C), annual precipitation mean 267 mm (n=23 years, all data from HYDROMET-CENTER ASTANA, 2000). The vegetation period is short, usually starting after snow melt and the last frosty nights in mid April, and lasting until end July.

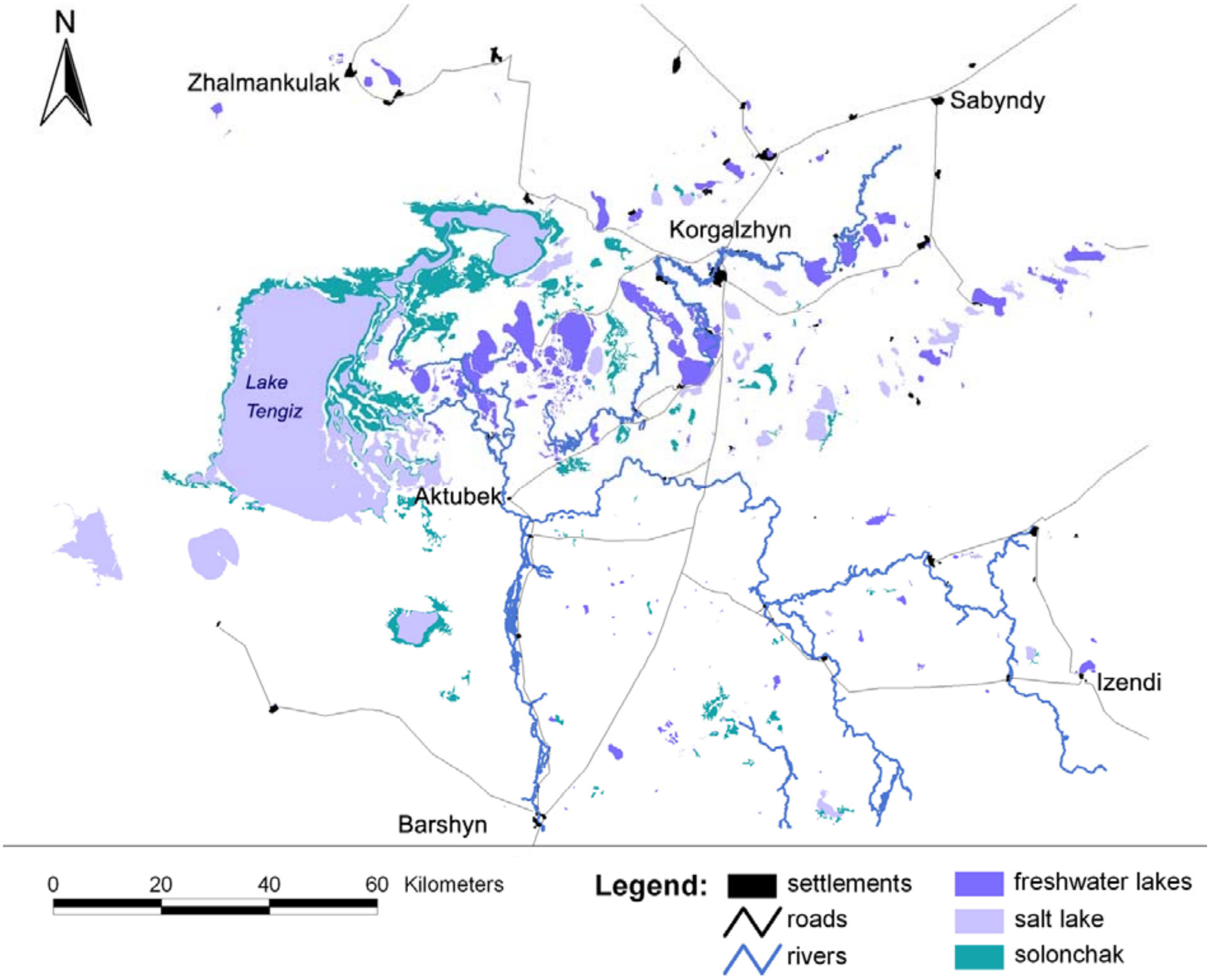


Fig. 3. Study area with main wetland complexes, rivers, roads and settlements.

Dominating soils within the study area are Kastanozems and Meadow-soils, the latter mostly along steppe rivers. About 30 % of the study area is characterized by Solonets and Solonchak salt soils. Solonets soils are often covered by short-grass steppe, whereas Solonchak areas appear often bare or sparsely vegetated. Nearly all mineral soils such as Kastanozems show traits of solonizing processes (PRASOLOVA and GERASIMOV 1955).

Soil type	**humus in upper layer (%)**	**thickness of the humus layer (cm)**
Chernozem	8-9	up to 90
Dark Kastanozem	6-7	50-70
Meadow soils	6-17	up to 100
Meadow-Kastanozem	<6	35-90
Pale Kastanozem	<2.5	30-45

Table 2. Main soil types in Korgalzhyn region, humus shares in the upper soil layer and mightiness of the humus layer (REDKOV 1964). Soil types names follow the Russian classification.

Kastanozems and Meadow-soils are characterized by mighty humus layers (Table 2), and humus accumulation is supported by bioturbation. Main rodent species responsible for this process are Steppe Marmot *Marmota bobac*, Little Souslik *Spermophilus pygmaeus*, Large-toothed Souslik *Spermophilus fulvus* and, to a minor extent, Five-toed Jerboa *Allactaga elater*.

The study area is situated in two ecoregions (biomes): “Kazakh steppe” and “Kazakh semi-desert” according to OLSON et al. (2001). Vegetation is characterized by Short-grass Steppe in the north, and Wormwood Semi-desert/Steppe in the south of the study area. In spring, steppe areas are covered with a carpet of geophytes such as several Tulip species (especially *Tulipa schrenkii*) (KARAMYSHEVA and RACHKOVSKAYA 1973). Semi-desert areas in the southern part of the study area are characterized by *Artemisietum* communities.

I distinguished between the following terrestrial “biotope types” applying an own vegetation classification based on KARAMYSHEVA and RACHKOVSKAYA (1973), since community-orientated classification systems by Soviet and Russian botanists (cf. KARAMYSHEVA and RACHKOVSKAYA 1973) are much too detailed for a bird habitat characterization. These biotope types were also used for a rough characterisation of Sociable Lapwing habitat use.

Short-grass Steppe Within the study area, this type occurs very often as "inlets" in larger *Stipa lessingiana* associations (transition to semi-desert), but becomes the dominant vegetation type where the landscape becomes more hilly. Short-grass Steppe is usually moderately to strongly grazed (Appendix D-1, D-3), only in remote areas ungrazed, and is dominated by some Chenopodiaceae and wormwood. Dominant plant species are *Festuca sulcata*, *Artemisia* spp., *Atriplex verruciferae* and *cana*, *Camphorosma* spec., *Galatella tatarica*. High cover of bare soil and gravel patches, often more than 50% of bare soil cover. Two variants, one very herb-rich (e.g. *Thymus* spec.), the other with high cover of mosses and lichen are found. If stronger grazed (especially around villages and cattle breeding brigades), *Artemisia schrenkiana*, *nitrosa* and other *Artemisia* species as well as *Taraxacum millefolium* reach massive dominance (*Artemisia* reaches up to 100% in total vegetation cover). In some areas, *Artemisia* is completely replaced by *Elymus ramosus* (Quack grass), probably caused by decreased grazing pressure.

Long-grass Steppe Near-natural *Stipa-Festuca* associations, with or without higher cover of *Spirea*-bushes, usually low grazing intensity (Appendix D-2). Grasses often show tussocky habitus, main grass species *Festuca sulcata*, *Stipa lessingiana* and *Stipa sareptana*.

Seed grass land Hay meadows, *Agropyron cristatum* very dominant (sown in until 1991), sometimes mixed with small shares of other grasses (*Elymus ramosus*, *Stipa lessingiana*, *sareptana* and *rubens*, *Festuca sulcata*), especially if fallow and not used for hay cutting for some years. Different stages of re-conversion into steppe communities.

Agricultural fields Fields with wheat (within the study area app.70% of all fields), barley (25 %) and oat (5%). Sawn in not earlier than mid May due to late frost in spring, harvested mostly in August.

Fallow fields Areas of former cereal and seed grass cultivation, fell fallow mostly since 1985, and again after the breakdown of the Soviet Union in 1991. Within the study area, approximately 70% of all earlier used fields are now fallow (own calculations). However, some of the fallow fields are ploughed up again in favourable years, but often fall fallow then after a period of two or three years use for cereal cultivation (Appendix D-4). Vegetation patterns on fallow fields are highly dependent on the intensity of former use, seed reservoirs and age (DIETERICH 2000 and pers. comm.). In the first year after abandonment, fields are often overgrown by single *Brassicaceae* species (especially *Brassica elongata*), evoking a yellow flowering aspect. Older stages become more “patchy”, with invading steppe herbs and a recolonisation by fescue and feather grass.

Typical species dominating patches are *Elymus ramosus*, *Sonchus arvensis*, *Petrosomia triandra* (in solonized areas), *Lactuca altaica*, *Ceratocarpus arenarius* and different *Artemisia* species, especially *Artemisia austriaca* (DIETERICH 2000). Fallow fields are used more and more for livestock grazing, which has a serious impact on species community composition and vegetation development. From observations in 2004 to 2006, I suspect grazing to have a considerable accelerating impact on re-conversion of fallow fields to Feather grass steppe.

Wet meadows In depressions where water accumulates during snow melt, wet, herb-rich meadow communities can be found. These are dominated by *Calamagrostis* grasses and herbs like *Salvia* spec.

Reeds In wet depressions and at shores of lakes and ponds, reeds of *Carex*, *Scirpus* and *Juncus* are found. *Phragmites* reeds grow at lakeshores and swamps where salt concentrations are not too high.

Solonchaks Salt pans and clayey salty depressions, at least partially flooded in spring (except after winters with very few snow), no or sparse vegetation cover with salt tolerating species. Usually dry, open soil in summer. Mostly bare, but partly covered by *Halocnemum strobilaceum*, *Salicornia* spp. and *Atriplex cana*. Not to be mixed up with Solonchak soil type.

Riparian shrub and bushes Up to 8 m height, dominated by Willow *Salix* spp. and *Rosa* spec. Found exclusively along steppe rivers.

Artificially planted shelter belts Up to 15 m height, wind shelter plantation along field edges and especially roads, comprising mainly introduced species such as Poplar *Populus spec.*, Maple *Acer negundo*, Elm *Ulmus minor* and *Eleagnos oxycarpa*.

2.2 Data sources and sampling strategy

2.2.1 Landscape basis data

To create a GIS basis for effective spatial analysis, Soviet topographic maps scaled 1:100,000 of the whole study area (last updated 1961-1989) were UTM geo-referenced and stored in ArcView 3.2a GIS. A geo-referenced Landsat 7 ETM+ satellite picture (custom-cut, resolution 14 meters) was saved as overlay to identify landscape changes after the map issue date and to characterise current land use. Based on this data, a digital map was created containing shape files with the main land use types and vegetation units, rivers, lakes and infrastructure. To gather additional information about the age of formerly agriculturally used fallow land, I checked land use maps of the territory of former state farms ("Sovkhoz") issued between 1960 and 1991 (ZELINGIPROZEM 1986, 1990a, 1990b). All spatial data collected (cf. chapter 2.2.2 – nest sites, rookeries) was transferred directly from the GPS receiver into ArcView GIS using the software Frontbase GPS 1.21.

2.2.2 Bird distribution and abundance

Surveys throughout the study area for Sociable Lapwing colonies were conducted between 25th April and 26th July 2006. Three teams of the Sociable Lapwing Project surveyed the study area spatially exhaustively. Survey trips were made daily starting half an hour before sunrise and ending at 8 p.m, with a break of usually one to two hours during the hottest period of the day. After findings of stationary and territorial birds, long-term observations from a distant vantage point or from a car used as a hide led to findings of incubating females. After spotting the location of the incubating female one person stayed on the spot watching carefully. This person then directed a second field worker to that spot using mobile radio units. This team-work of two persons was necessary, since breeding females leave the nest on approach of the investigator and fixing the point while walking proved to be difficult in many cases. Nests were marked by arranging some dry dung pieces 5 m off the nest site and coordinates of the nest where taken. The accuracy of GPS locations was 4±1.32 m (mean ± s.d., n=150 nests measured twice).

For the subsequent habitat variable recording, I considered colony and nest sites found until a cut off date of 25th May to avoid including colonies of breeders with failed first attempts and second clutches which might differ in selected habitat (e.g. BERG et al. 1992 for Northern Lapwing).

All sites within the study area containing higher shrub, bushes and planted woods including planted road shelterbelts were searched exhaustively for Rook (*Corvus frugilegus*) colonies between 1st June to 5th August, as Rook has been heavily suspected to be a major predator of Sociable Lapwing clutches (e.g. BELIK 2005). When a rookery was found, I recorded the number of occupied nests by passing through the colony on food and using mechanical hand counters. The number of occupied nests was considered equal to the number of 2006 breeding pairs (cf. Appendix C-1). Pair numbers at two colonies at inaccessible riparian shrub were estimated by dividing the maximum number of visible adults flying over the colony after disturbance by two.

2.2.3 Sampling design

I studied breeding habitat selection at the landscape scale and nest site selection within the colonies. As I decided to use presence-absence modelling without considering abundance, the data collected needed to be binary (cf. 2.3.1).

Sample unit at the **landscape scale** was the single colony. A colony was defined as a site with an accumulation (inter-nest spacing less than 500 meters) of -at least two- breeding pairs at least three kilometres apart from the next breeding incidence to guarantee statistical independence. The spatial extent of each colony was determined by buffering each nest site cumulatively with r=500m using the ArcView extension “X-Tools”. I chose these distance values according to my experience from 2004 and 2005 fieldwork periods that breeding adults during the incubation period rarely move more than 3-400 m away from the clutch and thus this distance should represent the area used around a selected nest site.
Within the colonies, variables were measured at a randomly chosen nest site representing “presence”. This was considered representative for the whole colony, since nest site characteristics at least for vegetation height and plant cover did not differ significantly within the single colony (tested for vegetation height and plant cover, “difference” was defined as total height/cover variation within the respective colony being larger than twofold standard deviation).

These points were compared with randomly selected points (representing “absence”) across the study area for analysis at the landscape scale. Random points were automatically defined on a map using the ArcView extension “RandomPoints” (JENNESS 2005) and were not allowed to fall within colony borders and water (exclusion with standard GIS tools). Coordinates from both presence and absence points were then transferred to GPS units and absence points could subsequently be easily identified in the field.

As logistical effort in point sampling increases strongly with increasing study area size, I decided not to sample across the whole study area, but only within a smaller study area of approximately 9,000 km² (Fig. 3). This area covers all habitat types on a gradient of increasing aridity to the south.

Effective sample size was limited by the scattered distribution of colonies throughout the study area. 30 colonies were available for variable sampling until the cut-off day of 25th May and habitat variables were recorded at 109 random points, thus prevalence was approximately 20% as suggested by BONN and SCHRÖDER (2001).

Sample unit at the **colony scale** was the single nest. At 17 colonies across the whole study area, I compared all nest sites with randomly chosen absence points within the borders of a colony (as defined above) characterizing habitat availability. Nests found later than May 25th were not considered as they might have concerned cases of re-nesting. Random points were not allowed to fall closer than 25 m to a nest site. The number of absence points per colony was determined to be approximately four times higher than the number of active nests, and thus prevalence was determined not to fall below 20 % as suggested by BONN and SCHRÖDER (2001). In total, I collected data at 78 nest sites and 262 randomly chosen absence points, resulting in a cumulated sample size of n=340 sample points, hence prevalence was 23 %.

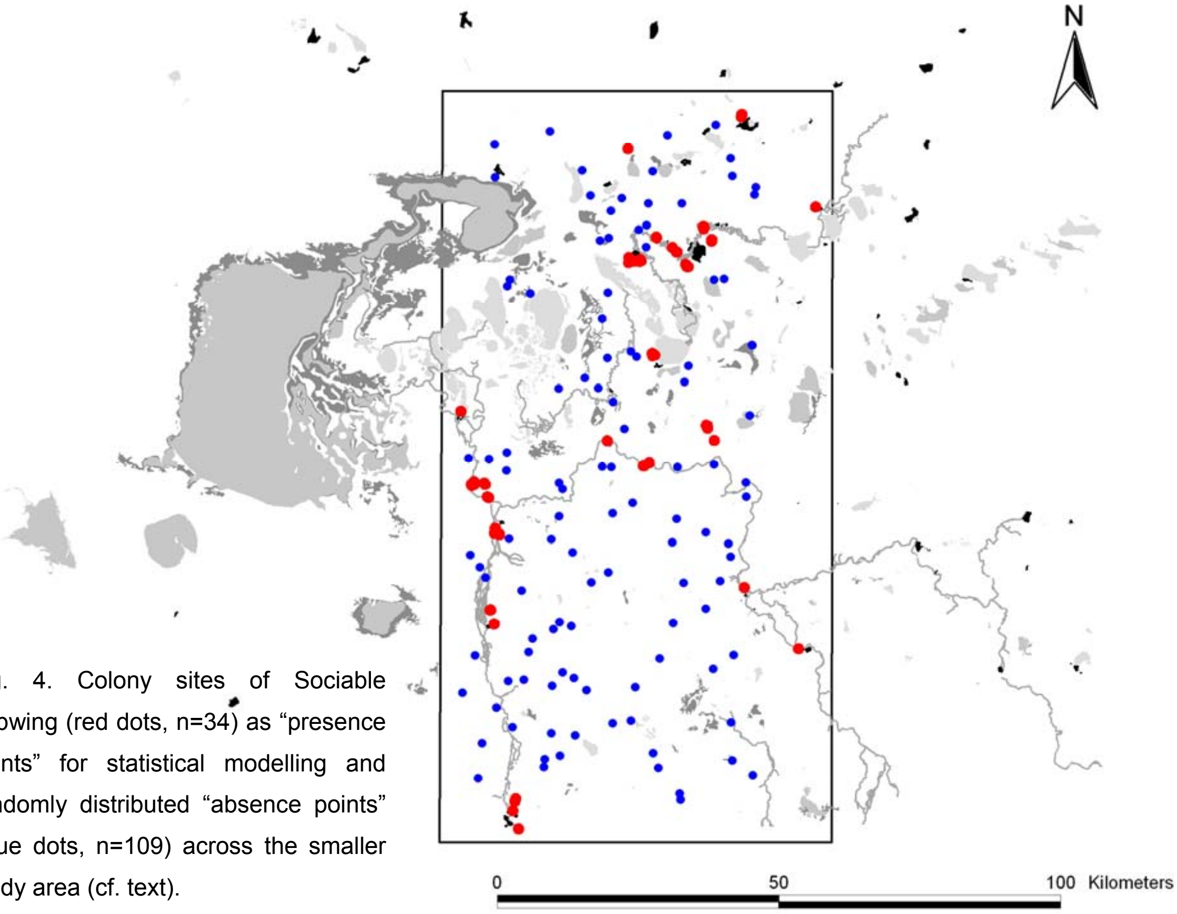


Fig. 4. Colony sites of Sociable Lapwing (red dots, n=34) as “presence points” for statistical modelling and randomly distributed “absence points” (blue dots, n=109) across the smaller study area (cf. text).

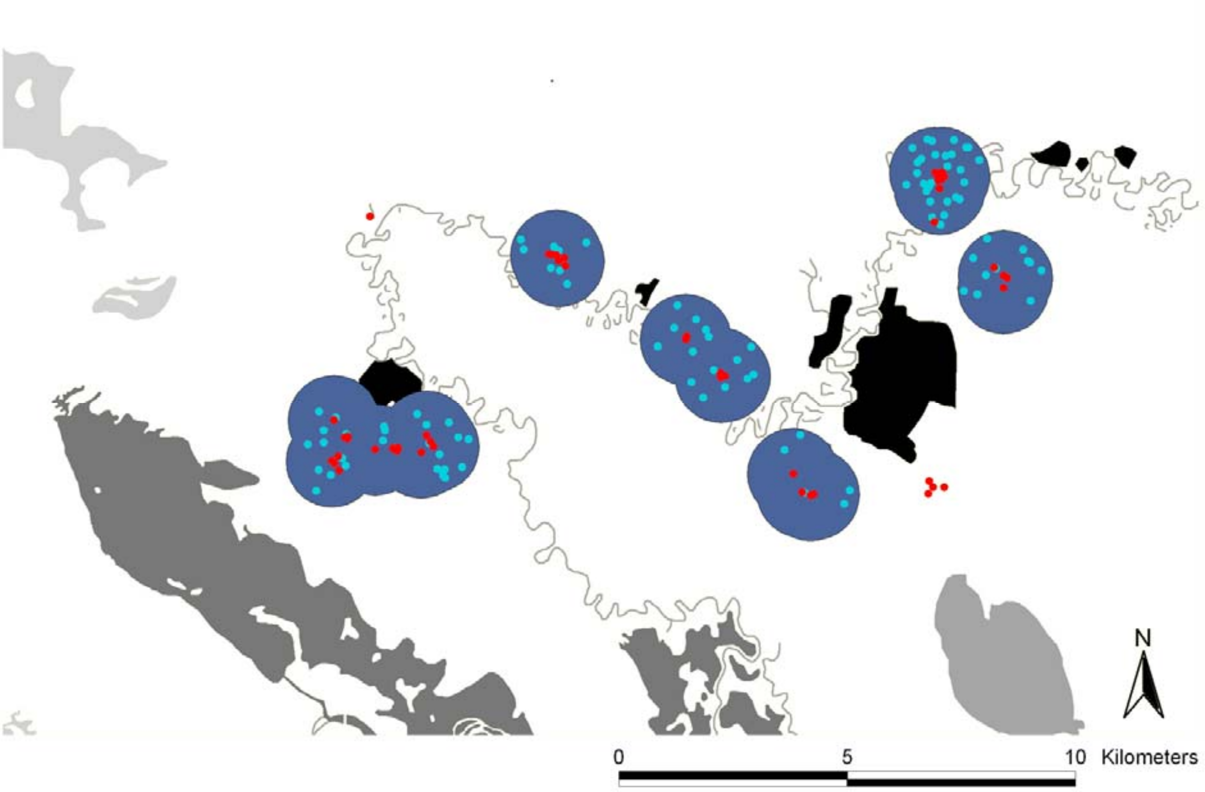


Fig. 5. Nest sites (red dots) of Sociable Lapwing as “presence points” for statistical modelling and randomly distributed “absence points” (cyan dots) within six “colonies” (blue buffer zones) around the village of Korgalzhyn. Nest sites outside colonies were all found after May 25th, assumed to be second clutches and not considered in analysis.

2.2.4 Variable recording

The choice of habitat parameters to be examined for an influence on Sociable Lapwing habitat selection was made hypothesis-based, with particular consideration of most of the literature available on Sociable Lapwing ecology and relying on own experience with the species from 2004 and 2005 (Tables 3,4). In total, I reviewed approximately 200 references containing data on Sociable Lapwing, mostly from Russian literature. Out of these, approximately 60 contained utilizable information on habitat use both in the breeding areas and during migration.

Table 3. Overview of all habitat parameters recorded on landscape level (for modelling habitat selection)

	variable	**shortcut**	**var_type**
topography	elevation above sea level (m)	elev	metric
	slope (°)	slop	metric
	aspect (transformed, y´=arccos (y) bzw. y´=arcsin (y))	asp	metric
hydrology	distance to nearest river (m)	dist.riv	metric
	distance to nearest standing water feature (m)	dist.lake	metric
	distance to nearest water feature (standing water or river, m)	dist.wat	metric
vegetation	cover of wormwood *Artemisia* spp. (%)	cov.art	metric
	cover of feather grass *Stipa* spp. (%)	cov.stip	metric
	cover of fescue *Festuca* spp. (%)	cov.fest	metric
	total grass cover (%)	cov.grass.tot	metric
	cover of herbaceous plants (without grasses) (%)	cov.herb	metric
	cover of mosses and lichens (%)	cov.ML	metric
	maximum vegetation height (mm)	vegH.max	metric
	modal vegetation height (mm)	vegH.mod	metric
soil	soil type (clay, silt, sand, gravel)	soil.type	kategoric
	soil surface structure (clay, silt, sand, gravel <10, 10-20, 20-50, >50 mm)	soil.surf	kategoric
stock farming	cover of dung (%)	cov.dung	metric
	dung abundance (fresh piles, absolute numbers)	dung.fresh	metric
	dung abundance (old piles, absolute numbers)	dung.old	metric
	total dung abundance (fresh and old piles, absolute numbers)	dung.tot	metric
public information and connectivity	distance to nearest Sociable Lapwing colony (m)	dist.col	metric
predation	distance to nearest rookery (m)	dist.rook	metric
disturbance	distance to nearest settlement (m)	dist.sett	metric

Table 4. Overview of all habitat parameters recorded on colony level (for modelling nest site selection)

	variable	**shortcut**	**var_type**
hydrology	distance to nearest river (m)	dist.riv	metric
	distance to nearest standing water feature (m)	dist.lake	metric
	distance to nearest water feature (m)	dist.wat	metric
vegetation	cover of wormwood *Artemisia* spp. (%)	cov.art	metric
	cover of feather grass *Stipa* spp. (%)	cov.stip	metric
	cover of fescue *Festuca* spp. (%)	cov.fest	metric
	total grass cover (%)	cov.grass.tot	metric
	cover of herbaceous plants (without grasses) (%)	cov.herb	metric
	cover of mosses and lichens (%)	cov.ML	metric
	maximum vegetation height (mm)	vegH.max	metric
	modal vegetation height (mm)	vegH.mod	metric
soil	soil type (clay, silt, sand, gravel)	soil.type	kategoric
	soil surface structure (clay, silt, sand, gravel <10, 10-20, 20-50, >50 mm)	soil.surf	kategoric
stock farming	cover of dung (%)	cov.dung	metric
public information and connectivity	distance to nearest Sociable Lapwing nest (m)	dist.nest	metric
predation	distance to nearest rookery (m)	dist.rook	metric
disturbance	distance to nearest settlement (m)	dist.sett	metric
life history needs	distance to significantly higher vegetation (m)	dist.Hveg	metric

Distance values were calculated using the extensions “nearest features”, “distance/azimuth tools” and “distance matrix” (JENNESS 2004, 2005a, 2005b) in ArcView 3.2a GIS. I calculated the distance to the nearest neighbouring colony (respectively nest) to estimate the influence of colonial breeding on habitat and nest site selection. The distance to nearest river was calculated, as there is evidence from recent colour ringing and satellite tagging that Sociable Lapwings migrate along steppe rivers as orientation strips in a flat and featureless landscape (R. SHELDON pers. comm.). This strategy might influence habitat selection at the landscape scale. The distance to the nearest standing water features was measured to estimate the importance of permanent water features in habitat selection (cf. WATSON et al. 2006), “distance to nearest water” was combined from the before mentioned.

A bird heavily suspected to play a special role in predation on Sociable lapwing clutches and chicks is the Rook *Corvus frugilegus* (BELIK 2005), thus I included distance to next rookery as variable in the models. The distance to settlements (villages and cattle breeding brigades) was located as an indicator for the intensity of disturbance: some authors evaluated disturbance as a critical factor in nest survival, thus strategies avoiding breeding too close to settlements might influence habitat selection (cf. WATSON et al. 2006, SOLOMATIN 1997). On the other hand, livestock is often concentrated around villages and might influence the distribution of Sociable Lapwing colonies (WATSON et al. 2006, SHELDON et al. in prep.).

The variable “distance to significantly higher vegetation” was measured only at colony level. It has been demonstrated before for Northern Lapwing, that birds select breeding sites with close-by higher vegetation that young chicks prefer for cover. “Higher vegetation” was defined as vegetation >20 cm taller than at the respective nest site/absence point. I chose this value relying on experience from the years before that females often guide their chicks into vegetation about this height after hatching (cf. also WATSON et al. 2006). I measured the distance directly in the field at every examined nest site resp. absence point walking the distance to the closest distinct taller vegetation and using the “go to” function of a Garmin GPS receiver.

Elevation, ***slope*** and ***aspect*** were recorded on landscape level, because some authors suspected micro-scaled differences in geomorphology as important in breeding site selection (e.g. preference of southern exposed microslopes, SOLOMATIN 1997). I copied elevation readings directly from the GPS unit, since test readings in Kazakhstan in advance showed that accuracy was not worse then 10m deviation in 95% of all cases (own tests in 2005, n=150 readings). Slope was measured with a clinometer and aspect was recorded with a simple compass (aspect reading accuracy =5°).

Repeatedly, it has been claimed that Sociable Lapwing prefers certain ***plant communities*** (KHROKOV 1977, BEREZOVIKOV et al. 1998, WATSON et al. 2006). To test this hypothesis, I recorded plant composition and coverage at 2x2 m sample plots, centred on the nest sites resp. absence points. Coverage was estimated as percentage precise to 5%, the level of estimate accuracy was checked by repeated sampling prior to the fieldwork period. As deviation in repeated samples was never more than 10% independent of observer, I regarded this method as adequate, however controlled every sampling myself during fieldwork. On a recording sheet prepared in advance, coverage of the most abundant steppe plant genera (Wormwood *Artemisia* spp., Fescue Grass *Festuca sulcata*, Feather Grass *Stipa lessingiana* and *sabulosa*) as well as total grass cover, cover of herbaceous plants (without grasses) and cover of mosses and lichens was recorded. Plant identification was done using SIDEROVA (1988) and SCHWAHN (2006).

Vegetation height was identified as an important factor in habitat selection of most open-country birds, especially waders *Charadriiformes* (CODY 1985, SMART et al. 2006). I measured vegetation height at every nest site and every absence point on two scales. I distinguished between “maximum vegetation height” which was defined as the height of the tallest plant(s) at the spot, and “modal vegetation height”, defined as the height of the majority of plants. The latter gives also an index for vegetation density (SUTHERLAND et al. 2004). Vegetation height was measured using a simple folding rule. Maximum vegetation height I read directly from the scale, whereas for modal vegetation height, I dropped a wooden disc (diameter 20 cm, height 0.4 cm, weight 80g) driving on the folding rule and read vegetation at the upper disc level, where it stopped (SUTHERLAND et al 2004). For each point I took five measurements in direct vicinity that were averaged.

The cover of bare ***soil*** was estimated analogue to the plant coverage estimates (see above). Additionally, I reported soil type and soil surface structure, using the classes clay, sand, silt, gravel >10mm, gravel 10-20mm, gravel 20-50 mm, gravel >50mm grain size/diameter. Bare rock as well as tarmac or other artificial surface structure was grouped into the latter category. All classes were later transformed into a metrical scale from 1-7, since they represent a gradual transition in grain size.

Sociable Lapwing probably co-evolved with steppe ungulates and is thus dependent on the presence of large grazing animals (WATSON et al. 2006, SOLOMATIN 1997, KHROKOV 1977). Nowadays, wild ungulates such as Saiga antelope *Saiga tatarica* or Kulan wild ass *Equus hemionus* have disappeared nearly entirely from the steppes west of the Altai (BEKENOV et al. 1998, MILNER-GULLAND et al. 2001, FEH et al. 2002), but have been replaced by domestic livestock (ROBINSON and MILNER-GULLAND 2003). In consequence, I tried to estimate the dependence of Sociable Lapwing habitat selection from the presence and abundance of grazing livestock. Since I was not able to estimate absolute livestock numbers spatially exhaustive, and as ***dung*** is a very good correlate of grazer density (LAING et al. 2003), I characterized every examined nest site and every absence point by the cover of dung analogue to the plant coverage estimates (see above). This proved an ineffective method for larger scales, since livestock is comparatively mobile and dung density low, so on landscape scale I additionally counted all livestock dung piles separated for sheep, horses and cattle, and separated for fresh and old piles over a transect of 25 m both to the west and east of each nest site respectively absence point (BUCKLAND et al. 2001).

Fieldwork on parameter recording was conducted between May 16th and May 25th 2006 at the landscape scale, respectively between May 25th and June 5th 2006 within the colonies.

To get a first overview of spatial movements of livestock flocks, two Garmin GPS units were fixed on two cows in different cattle herds at the village of Korgalzhyn using simple dog collars. The receivers were connected with an automatic timer, programmed to switch the device on every morning at 6 a.m. and off at 7 p.m.. Using the “track” option, the GPS receiver was programmed to save the cow’s (and hence the flock’s) position every two minutes. Due to technical problems, we were able to collect data only on five days in May, not enough to allow analysis.

2.3 Data Analysis

2.3.1 Model building

As modelling strategy I chose logistic regression on both landscape and colony level, since this presence-absence method proved to be relatively robust concerning deviation from normal distribution (BACKHAUS et al. 2003), and gives much better results compared to presence only models (BROTONS et al. 2004, ENGLER et al. 2004). Logistic regression is a special case of Generalized Linear Modelling (GLM) with a logistic link function, for details cf. HOSMER and LEMESHOW (2000).

To select parsimonious models with high information content, I proceeded as follows. I first started with a careful univariate analysis to eliminate variables not correlated with presence or absence and thus avoided spurious inclusion of variables in multiple models as suggested by HOSMER and LEMESHOW (2000). I modelled all variables applying logistic regression using the “logistf” package for S-Plus 6.1 by HEINZE and SCHEMPER (2002). According to MANLY (2001), I bootstrapped each univariate model 300 times. After each bootstrap, deviance reduction was recorded and a Likelihood-Ratio (LR-) test conducted. I included variables into the further multivariate modelling process, if the bootstrap-LR-test was significant ($p \leq 0.05$) for at least 95% of the bootstrap iterations and thus a significant correlation between variable characteristic and presence/absence was suggested. At this stage, some strongly correlated (Spearman’s rank coefficient r_S, two-sided test, as normal distribution and linearity not guaranteed) parameters describing the same feature (e.g. dung coverage vs. dung transect abundance) were excluded from the further procedure. Out of a set of correlated variables, I chose the one with the strongest relationship, usually accompanied by highest R^2_N (NAGELKERKE 1991) and AUC (HANLEY 1982) as measures of model calibration and refinement respectively discriminative power. Correlation was assumed if $r_S \geq 0.5$. The latter criteria were chosen to early eliminate variables, which proved significant, but had limited predictive power and thus would hardly improve multivariate models but rather exaggerate further computing processes.

Before processing variables in multiple models, I determined the shape of the environmental relationship, i.e. the pattern of each variable’s response to environmental factors. This has often been neglected, but is a crucial factor when using statistic models to predict species abundance and distribution (AUSTIN 2002). Relationships can be either sigmoid (with positive or negative characteristic) or unimodal (“bell-shaped” or “bowl-shaped”). If a relationship was sigmoid or unimodal, I decided by the level of significance. If both relationships were significant, I chose the one with the stronger response (STRAUSS and BIEDERMANN 2006). I finally plotted response curves for every univariate model to control for ecologically nonsense relationships.

For multivariate modelling, I applied multivariate logistic regression at both landscape and colony scale. I used a backward stepwise approach for model simplification, which is still widespread among ecologists (e.g. HARREL 2001, OPPEL 2004, LÓPEZ-LÓPEZ 2005). I first decreased multicollinearity by removing out of a pair of two variables the one with the lower Spearman's r (for correlation matrices cf. App. A-1, A-2). Then, variables were removed until further removal would have led to a significantly worse model indicated by significant deviance reduction (HOSMER and LEMESHOW 2000). The obtained model was considered the "best model". In each step, the critical p-value was $p_{out}=0.10$ for variable removal. I used p-values from LR test statistics, since Wald test has been criticised massively in recent publications (HEINZE and SCHEMPER 2002, BURNHAM and ANDERSON 2002). All stepwise procedures were calculated in SPSS 13.0.

However, stepwise significance testing faced increasing, recently strengthened criticism (ANDERSON et al. 2000, GREAVES et al. 2006, POSADA and BUCKLEY 2004, WHITTINGHAM et al. 2006). This method has been criticized mainly because of i) the use of ecologically irrelevant null-hypotheses, ii) arbitrarily chosen significance-levels and iii) the neglect in consideration of ecological background information during the stepwise computing processes (BURNHAM and ANDERSON 2002). The "best-fit model" obtained by this method is often accompanied by several models only marginally worse, which are not considered then.

As a currently establishing alternative, I chose an information-theoretic approach comparing all possible multivariate models and comprising model averaging (BURNHAM and ANDERSON 2002, STRAUSS and BIEDERMANN 2006, GREAVES et al. 2006). With an automated procedure, I calculated models for all possible combinations of four, three and two variables using a self-programmed script for S-Plus 6.1 (Appendix B-1). As sample size was comparatively low in this study (cf. 2.2), including more than four variables in the same model would have led to over-parametrisation in multivariate models (REINEKING and SCHRÖDER 2004, GUISAN and ZIMMERMANN 2000). Furthermore, models including very many variables are difficult to interpret. As including correlated variables into the same model can lead to severe bias (problem of multicollinearity, GRAHAM 2003), I allowed only combinations of variables with $r_S<0.5$ to appear in the same model. FIELDING and HAWORTH (1995) and OPPEL et al. (2004) allowed a threshold of $r_S=0.7$, but given that multicollinearity decreases statistical power often at a level of $r_S=0.3$ already (GRAHAM 2003), I decided to be more strict. As I intended to achieve parsimonious models with high information power, I conducted a LR-test for every model to test if it was better than (or just as good as) any model with one variable less (FERRIER et al. 2002). Additionally, I tested, whether corrected R^2_N (200 bootstrapping iterations) was ≥0.3 (STRAUSS and BIEDERMANN 2006). If both requirements were fulfilled, the achieved model was considered "adequate". Univariate models with a R^2_N ≥0.4 (model calibration and refinement "good" according to BACKHAUS et al. 2003) were considered as "multivariate models with one variable" and regarded adequate as well.

2.3.2 Model selection and model averaging

As on both landscape and colony level many adequate models were achieved, I proceeded using the model selection approach developed by BURNHAM and ANDERSON (2002). For each model, the AIC ("An Information Criterion", often referred to as "Akaike Information Criterion", AKAIKE 1973, ANDERSON et al. 2000) was calculated, a measure to describe how well models perform the trade-off between model fit and model complexity:

$$AIC = -2LL + 2p \tag{1}$$

$$AIC_C = \frac{AIC + 2p(p+1)}{n - p - 1} \tag{2}$$

where LL=LogLikelihood of the model, p=number of regression coefficients + number of constants + number of variance parameters, n=covariate sample size.

I used the corrected value AIC_C (term 2), as suggested by BUCKLAND et al. (1997) when sample size is small ($n/p<40$). The model with the lowest AIC_C has the best relative fit (within the considered set of models), given the number of parameters included. For each model, Akaike weights w_i of each model i were calculated (term 4), using the AIC_C differences Δ_i (term 3) between the best model and each other model:

$$\Delta_i = AIC_{C_i} - AIC_{C\,\min} \tag{3}$$

$$w_i = \frac{\exp(-0.5\Delta_i)}{\sum_{m=1}^{M} \exp(-0.5\Delta_m)} \tag{4}$$

Akaike weights are based on Kullback-Leibler information in the sense of the existence of a "best model" regarding full reality, data, sample size and model set. A given w_i is considered as the weight of evidence in favour of model i being the actual Kullback-Leibler best model for the situation at hand - given that one of the M models must be the Kullback-Leibler best model of a set of M models (BURNHAM and ANDERSON 2002).

According to BURNHAM and ANDERSON (2002) models with $\Delta_i<2$ have the highest level of empirical support of model i. Models with Δ_i from 4 to 7 are classified as having a considerably less, models with Δ_i >10 have either no empirical support of model i, or at least fail to explain some substantial variation in the data. Furthermore, if the weights of all models are summed up cumulatively (starting at the best model), the 0.95 share can be used as an equivalent of 95% certainty.

Using Akaike weights, for both landscape and colony level, the Kullback-Leibler “best models”, i.e. those having the highest probability to be closest to reality, were determined. I used those having $\Delta_i<10$, up to a cumulated share of 0.95 (GREAVES 2006).

Finally, I applied model averaging for all models as suggested by BURNHAM and ANDERSON 2002) using a self-programmed script in S-Plus (Appendix B-2). This has proved a suitable method to analyse data sets with variables of similar predictive power (e.g. STRAUSS and BIEDERMANN 2006). Model coefficients were weighted with the weight of every model they appeared in. The sum over all models for all weighted coefficients for a given variable represents the averaged coefficient for this variable. By this, the sought-after averaged model was obtained.

2.3.3 Internal model validation

To assess the predictive power of the achieved models, I calculated a set of criteria describing model fit (calibration and refinement) and discriminative power. This is necessary, since there is no criterion which can be used to describe all aspects of model quality together.

In detail, I calculated i) Nagelkerke’s R^2, which describes model calibration and refinement (NAGELKERKE 1991), ii) AUC (area under the receiver operating characteristic), which describes discrimination (HANLEY 1982, MCNEIL and HANLEY 1984), sensitivity (share of correctly predicted species’ presences), specificity (share of correctly predicted species’ absences) and CCR (overall correct classification rate) for discriminative power (FIELDING and BELL 1997, HOSMER and LEMESHOW 2000), and iv) Cohen’s Kappa, which describes discriminative power prevalence-independent (MANEL et al. 2001).

All models were internally validated using bootstrapping (MOONEY and DUVAL 1993, MANLY 2001) with 300 iterations.

2.3.4 Predictive modelling and external model validation

Predictions on species occurrence in landscapes can be done applying the regression equation to spatial grid data (OSBORNE et al. 2001, AUSTIN 2002, JOY and DEATH 2004). This process requires exhaustive coverage for the parameters (covariates) included into the model referred to. Out of all models containing only variables exhaustively covered for the whole study area (landscape scale only) I chose the best, i.e. the model with lowest Δ_i, and used it for spatial modelling.

I created grid themes in ArcView 3.2a describing the value of the considered variables for every grid cell. Grid resolution was determined to be b=50 m. I applied the logistic regression equation for the referring model to each grid cell using the “Map Calculator” included in the ArcView standard extension “Spatial Analyst”. Via a classifying process, areas of the same occurrence probability were ranked equally. Furthermore, I used a second classification to determine “suitable habitat”. I chose p_{Kappa}, i.e. the p-value where Kappa is maximised, as a threshold distinguishing between "suitable" and "not suitable". This threshold has proved to distinguish most precisely among different p-values. Finally, the analysis results were plotted to habitat suitability maps.

To externally validate the model on a temporal scale, i.e. to test temporal model generality, I determined to which extent nest sites from previous years (mapped in 2004 and 2005) were situated within the areas labelled as "suitable".

3 RESULTS

3.1 Habitat use

In total, 168 nests were found in 34 colonies (Fig. 12, p. 33). Within the study area, Sociable Lapwing used predominantly intensively grazed areas. Colonies were found within the following main biotope types (cf. 2.1): Short-grass Steppe, fallow wheat and seed grass communities usually dominated by Wormwood *Artemisia*, recently (autumn 2005 and spring 2006) burnt steppe (predominantly Feather Grass *Stipa*-communities), and dry Solonchaks with dense *Salicornia* cover (Fig. 6).

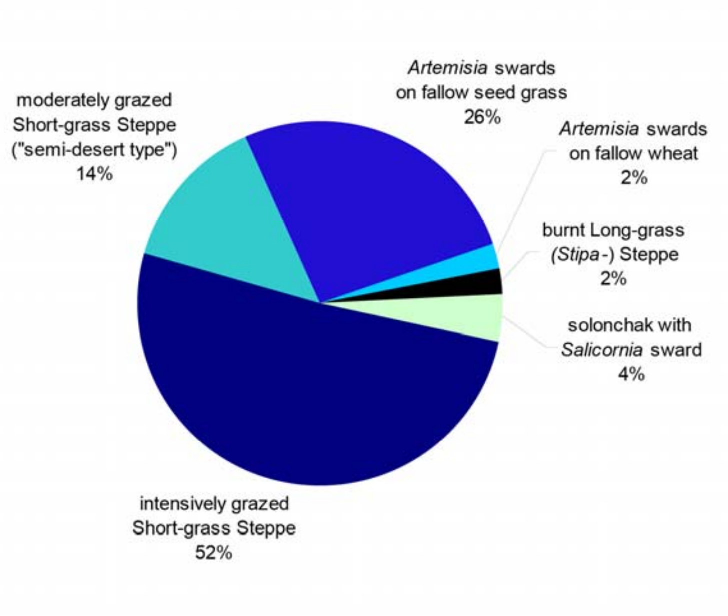


Fig. 6. Habitat use of Sociable Lapwing in Korgalzhyn region 2006 (main habitat types, n=168 nests).

3.2 Habitat selection: Univariate models

3.2.1 Landscape scale

At the landscape scale (“breeding habitat selection”), 16 variables showed significant ($p<0.05$) correlations between species occurrence and parameter characteristics in univariate binary logistic regression models. 13 relations were significant at a level of $p<0.001$ (Table 5, cf. also App. A-3).

Table 5. p-values (LR-Test) and model performance (R^2_N and AUC) of all variables in univariate logistic regression models, landscape scale.

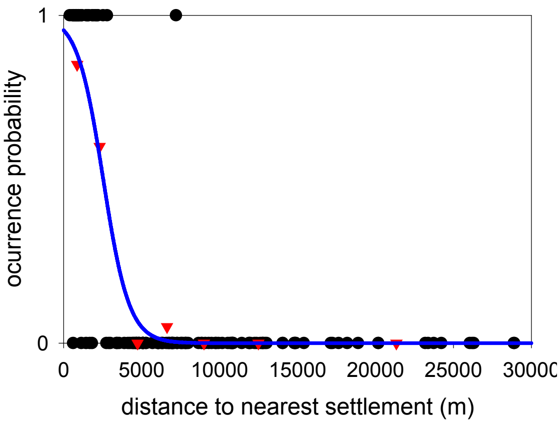
variable	**shortcut**	**p sigm**	**p unimod**	**R^2_N sigm**	**AUC sigm**	**R^2_N unimod**	**AUC unimod**
distance to nearest settlement (m)	dist.sett	0.0000	0.0000	0.759	0.964	0.761	0.963
total dung abundance (fresh and old piles, absolute numbers)	dung.tot	0.0000	0.0000	0.569	0.929	0.592	0.928
dung abundance (old piles, absolute numbers)	dung.old	0.0000	0.0000	0.524	0.917	0.542	0.917
modal vegetation height (mm)	vegH.mod	0.0000	0.0000	0.448	0.896	0.436	0.895
maximum vegetation height (mm)	vegH.max	0.0000	0.0000	0.417	0.869	0.397	0.867
distance to nearest river (m)	dist.riv	0.0000	0.0000	0.392	0.869	0.435	0.865
distance to nearest standing water feature (m)	dist.wat	0.0000	0.0000	0.384	0.831	0.364	0.831
dung abundance (fresh piles, absolute numbers)	dung.fre	0.0000	0.0000	0.291	0.808	0.311	0.807
cover of dung (%)	cov.dung	0.0000	0.0000	0.286	0.875	0.315	0.874
cover of Feather Grass *Stipa* spp. (%)	cov.stip	0.0000	0.0000	0.231	0.716	0.224	0.715

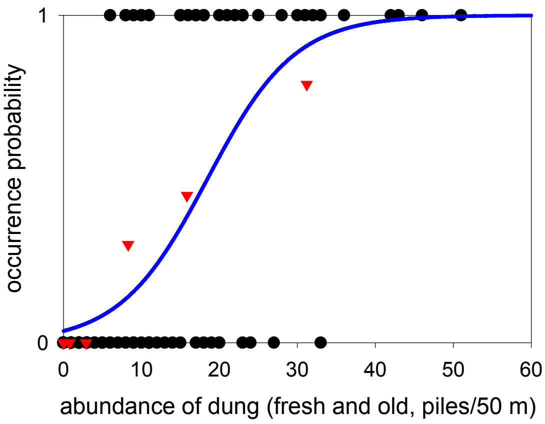
total grass cover (%)	cov.gras.tot	0.0000	0.0000	0.225	0.738	0.247	0.739
cover of herbaceous plants (without grasses) (%)	cov.herb	0.0000	0.0001	0.193	0.740	0.188	0.737
cover of Wormwood *Artemisia* spp. (%)	cov.art	0.0003	0.0000	0.138	0.761	0.259	0.786
distance to nearest rookery (m)	dist.rook	0.0010	0.0002	0.115	0.713	0.175	0.674
total vegetation cover (%)	cov.veg	0.0022	0.0000	0.101	0.739	0.252	0.768
distance to nearest Sociable Lapwing colony (m)	dist.col	0.0128	0.0060	0.067	0.673	0.109	0.645
slope (°)	slope	0.0858	0.3083	0.032	0.577	0.026	0.571
cover of Fescue Grass *Festuca* spp. (%)	cov.fest	0.0873	0.2630	0.032	0.582	0.029	0.581
aspect (arc/cos-transformed)	asp	0.0952	0.2533	0.031	0.574	0.030	0.574
elevation above sea level (m)	elev	0.4838	0.7232	0.005	0.537	0.007	0.551
cover of mosses and lichens (%)	cov.ML	0.5224	0.8160	0.005	0.543	0.005	0.543
distance to nearest water feature (m)	dist.lake	0.9399	0.9691	0.000	0.512	0.001	0.549
soil surface structure (clay, silt, sand, gravel <20, 10-20, 20-50, >50 mm)	soil.surf	0.9510	n/a	n/a	n/a	n/a	n/a
soil type (clay, silt, sand, gravel)	soil.type	1.0000	n/a	n/a	n/a	n/a	n/a

After the following process of variable selection (cf. 2.2) with the removal of all non-significant relationships (model simplification), 10 variables remained to be included in multivariate models (Table 6). The variables “cover of bare soil”, “distance to nearest rookery” and “distance to nearest Sociable Lapwing colony” did not pass the bootstrapping procedure, “distance to nearest Sociable Lapwing colony” additionally had very low R^2_N. The variables “cover of dung” and “dung abundance fresh/old” were strongly correlated with “total dung abundance” ($r_S>0.7$). “Maximum vegetation height” was strongly correlated with “modal vegetation height” ($r_S=0.83$). Six relationships were significantly sigmoid negative, one sigmoid positive, and three unimodal (Table 6).

Table 6: Variables selected for multivariate analysis and the according p-values, R^2_N, AUC and direction of environmental response, landscape scale.

variable	shortcut	p	R^2_N	AUC	Response
distance to nearest settlement (m)	dist.sett	0.0000	0.759	0.964	-
total dung abundance (fresh and old piles, absolute numbers)	dung.tot	0.0000	0.569	0.929	+
modal vegetation height (mm)	vegH.mod	0.0000	0.448	0.896	-
distance to nearest river (m)	dist.riv	0.0000	0.392	0.869	-
distance to nearest standing water feature (m)	dist.wat	0.0000	0.384	0.831	-
cover of Feather Grass *Stipa* spp. (%)	cov.stip	0.0000	0.231	0.716	-
total grass cover (%)	cov.gras.tot	0.0000	0.247	0.739	+/- unimodal
cover of herbaceous plants (without grasses) (%)	cov.herb	0.0000	0.193	0.740	-
cover of Wormwood *Artemisia* spp. (%)	cov.art	0.0000	0.259	0.786	+/- unimodal
total vegetation cover (%)	cov.veg	0.0000	0.252	0.768	+/- unimodal





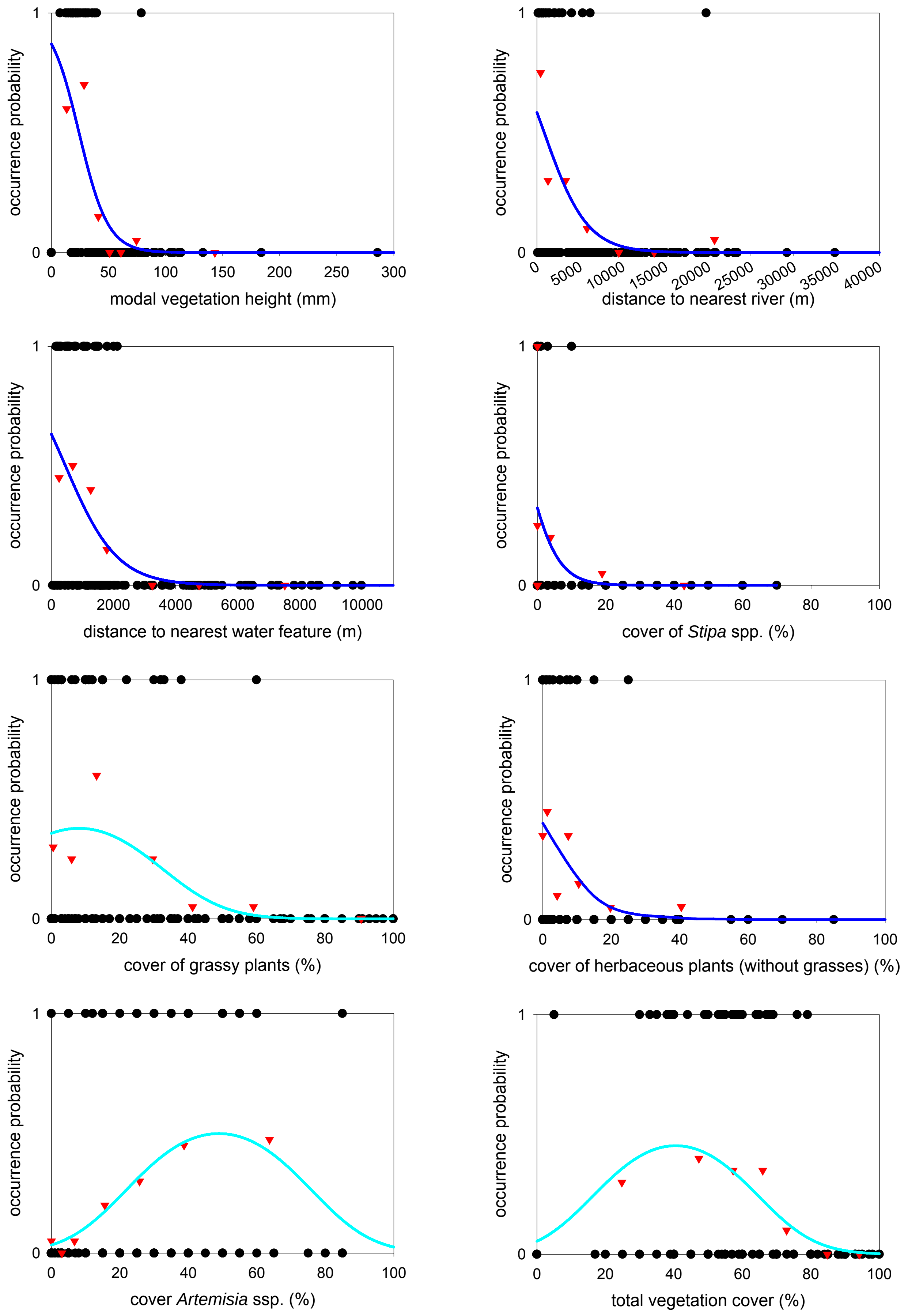


Fig. 7. Response shapes for all variables considered in multivariate analysis (landscape scale). A blue graph indicates sigmoid relationships, a cyan unimodal ones. ● = presence/absence data, ▲ = octiles of the data set.

3.2.2 Colony scale

At the colony scale ("nest site selection"), eight variables showed significant ($p<0.05$) correlations between species occurrence and parameter characteristics in univariate binary logistic regression models. Five relations were significant at a level of $p<0.001$ (Table 7, cf. also App. A-4).

Table 7. p-values (LR-Test) and model performance (R^2_N and AUC) of all variables in univariate logistic regression models, colony scale.

variable	shortcut	p sigm	p unimod	R^2_N sigm	AUC sigm	R^2_N unimod	AUC unimod
distance to nearest Sociable Lapwing nest (m)	dist.nest	0.0000	0.0000	0.460	0.895	0.600	0.927
maximum vegetation height (mm)	vegH.max	0.0000	0.0000	0.172	0.739	0.169	0.738
modal vegetation height (mm)	vegH.mod	0.0000	0.0000	0.132	0.701	0.129	0.700
cover of dung (%)	cov.dung	0.0000	0.0000	0.131	0.723	0.175	0.728
total grass cover (%)	cov.grass	0.0000	0.0000	0.092	0.596	0.118	0.605
distance to significantly higher vegetation (m)	dist.Hveg	0.0069	0.0016	0.042	0.647	0.073	0.657
total vegetation cover (%)	cov.veg	0.0073	0.0000	0.032	0.617	0.107	0.671
cover of Wormwood *Artemisia* spp. (%)	cov.Art	0.0174	0.0041	0.025	0.606	0.048	0.608
soil type (clay, silt, sand, gravel)	soil.type	0.0669	0.0597	0.016	0.533	0.027	0.587
soil surface structure (clay, silt, sand, gravel <20, 10-20, 20-50, >50 mm)	soil.surf	0.0669	0.0597	0.016	0.533	0.027	0.587
cover of Feather Grass *Stipa* spp. (%)	cov.Stip	0.0918	0.3414	0.013	0.517	0.010	0.517
cover of mosses and lichens (%)	cov.ML	0.1565	0.1675	0.009	0.595	0.016	0.596
distance to nearest rookery (m)	dist.rook	0.1723	0.2195	0.008	0.525	0.013	0.539
distance to nearest settlement (m)	dist.sett	0.3343	0.5341	0.004	0.547	0.006	0.547
cover of Fescue Grass *Festuca* spp. (%)	cov.Fest	0.5990	0.1566	0.001	0.484	0.016	0.567
cover of herbaceous plants (without grasses) (%)	cov.herb	0.6368	0.8824	0.001	0.538	0.001	0.473
distance to nearest standing water feature (m)	dist.lake	0.6602	0.8066	0.001	0.511	0.002	0.507
distance to nearest water feature (m)	dist.water	0.7111	0.3375	0.001	0.536	0.010	0.541
distance to nearest river (m)	dist.riv	0.8152	0.9698	0.000	0.479	0.000	0.531

After the following process of variable selection (cf. 2.2) with the removal of all non-significant relationships (model simplification), five variables remained to be included in multivariate models (Table 8). The variables "modal vegetation height", "distance to significantly higher vegetation" and "cover of Wormwood *Artemisia* spp." did not pass the bootstrapping procedure.
Three relationships were significantly negative and two unimodal (Table 8).

Table 8. Variables selected for multivariate analysis and the according p-values, R^2_N, AUC and direction of environmental response, landscape scale.

variable	shortcut	p	R^2_N	AUC	Response
distance to nearest Sociable Lapwing nest (m)	dist.nest	0.0000	0.460	0.895	-
maximum vegetation height (mm)	vegH.max	0.0000	0.172	0.739	-
cover of dung (%)	cov.dung	0.0000	0.175	0.728	+/- unimodal
total grass cover (%)	cov.grass	0.0000	0.092	0.596	-
total vegetation cover (%)	cov.veg	0.0000	0.107	0.671	+/- unimodal

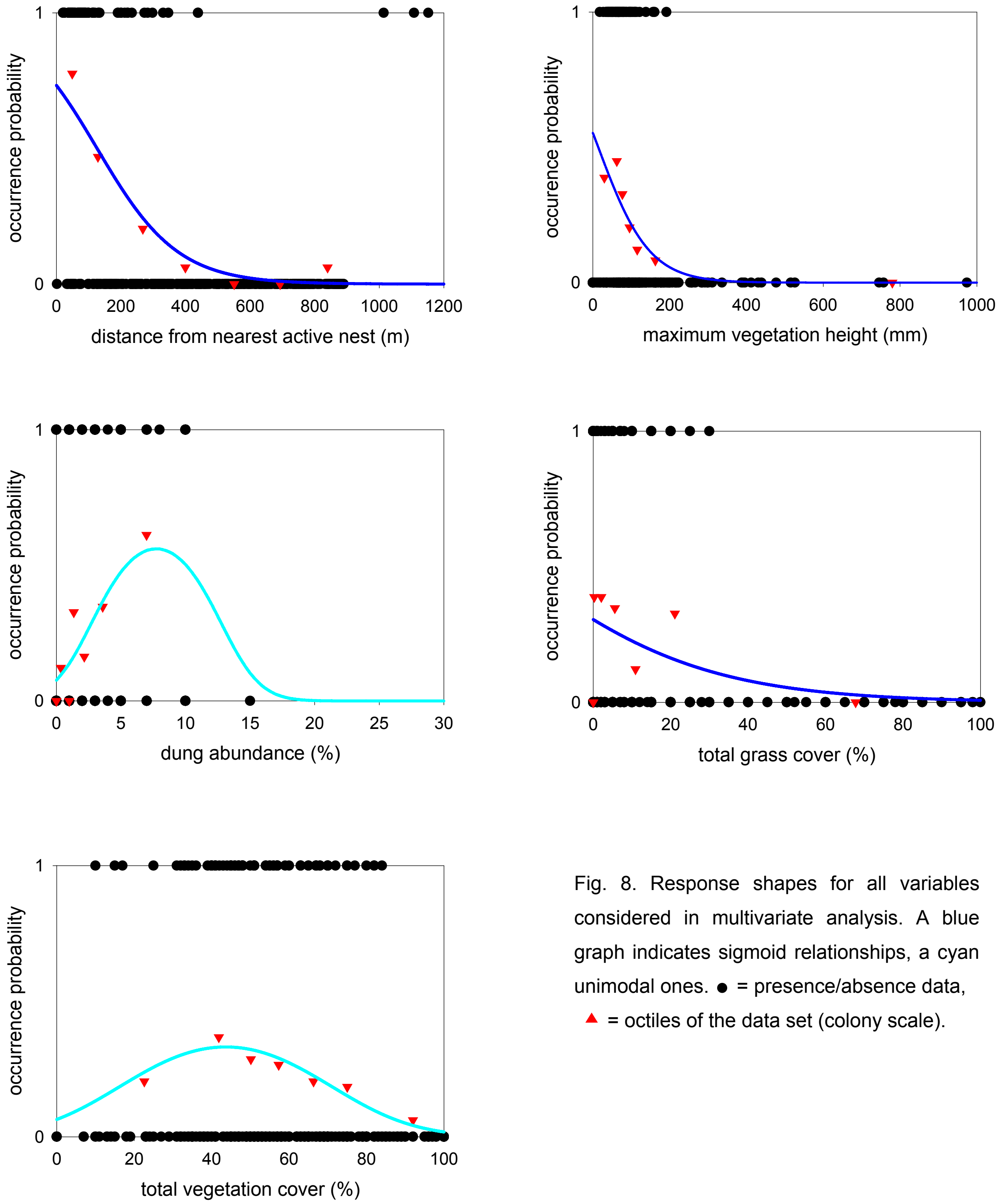


Fig. 8. Response shapes for all variables considered in multivariate analysis. A blue graph indicates sigmoid relationships, a cyan unimodal ones. ● = presence/absence data, ▲ = octiles of the data set (colony scale).

3.3 Habitat selection: Multivariate models

3.3.1 Landscape scale

The best model at the landscape scale (“breeding habitat selection”) achieved by logistic regression with stepwise significance testing identified “distance to nearest settlement”, “maximum vegetation height” and “distance to river” as the most influent parameters (logit: P(y=1) = 5.500 –

0.008 x dist.sett + 0.0143 x vegH.max – 0.002 x vegH.max^2 – 0.002 x dist.riv). The model correctly classified 98.6% of the data points, and internal validation did not significantly reduce model quality. Overall model performance was excellent (Nagelkerke's R^2=0.916, $AUC_{bootstrapped}$=0.995, Cohen's κ=0.897)

Within the alternative modelling process, at the landscape scale, 61 models qualifying as "adequate" were obtained (Appendix A-5), including three models with only one variable.

In the following weighting process, ten models remained with w_i different from zero (Table 9). Omitting those with no empirical support (Δ_i>10), at the landscape scale five models with acceptable fit remained (Table 9). The weights of already the first three models summed up to 0.96.

Table 9. Models with w_i>0, their AIC_C differences to the model with the highest AIC_C (Δ_i) and their empirical support of model i according to BURNHAM and ANDERSON (2002), landscape scale.

ID	models	AICc	Δ_i	w_i	level of empirical support of model i
1	vegH.mod +dist.riv +dist.sett	36.441	0.000	0.679	substantial
2	cov.art +cov.art^2 +cov.gras.tot +cov.gras.tot^2 +dist.wat +dung.tot	38.474	2.033	0.246	substantial
3	cov.art +cov.art^2 +vegH.mod +dist.riv +dist.wat	42.363	5.921	0.035	considerably less
4	vegH.mod +dist.riv +dist.wat +dung.tot	43.013	6.572	0.025	considerably less
5	cov.stip +dist.riv +dist.sett	46.281	9.840	0.005	essentially none
6	vegH.mod +dist.sett	47.437	10.996	0.003	essentially none
7	cov.art +cov.art^2 +cov.stip +dist.wat +dung.tot	48.179	11.738	0.002	essentially none
8	cov.art +cov.art^2 +dist.riv +dist.wat +dung.tot	48.260	11.818	0.002	essentially none
9	vegH.mod +dist.riv +dung.tot	48.276	11.835	0.002	essentially none
10	dist.riv +dist.sett	49.799	13.358	0.001	essentially none

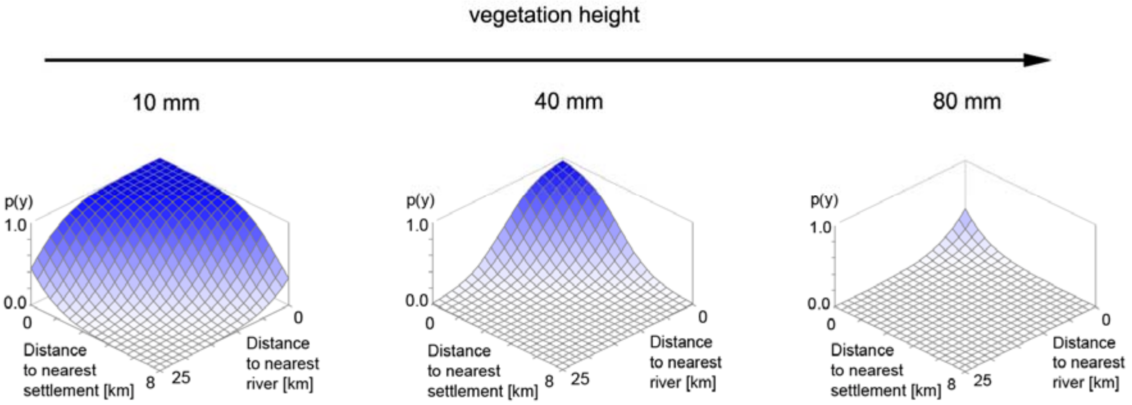


Fig. 9. Habitat model with lowest AIC_C ("best model") at landscape level. P is plotted against "distance to nearest settlement" and "distance to nearest river". Diagrams represent different stages of modal vegetation height [AUC=0.99, R^2_N=0.89, Cohen's κ=0.96, all values bootstrapped].

Model averaging resulted in three variables with very high weights, four with medium and three with very low weights (Table 10.).

Table 10. Averaged coefficients, variable weights and variable ranking for the landscape scale model set.

variable	averaged β	weight	rank
Intercept	3.977	1.000	-
dist.riv	-0.209	0.749	1
vegH.mod	-83.884	0.744	2
dist.sett	-0.718	0.688	3
dist.wat	-1.242	0.311	4
cov.art	0.242	0.285	5
cov.art^2	-0.003	0.285	
dung.tot	0.227	0.277	6
cov.gras.tot	-0.237	0.246	7
cov.gras.tot^2	0.002	0.246	
cov.stip	-0.064	0.007	8
cov.veg	0.147	0.000	9
cov.veg^2	-0.002	0.000	
cov.herb	-0.087	0.000	10

According to both Δ_i values of the referring models and variable ranking due to AIC weights (Table 10), it became clearly obvious that the variables "distance to nearest river", "distance to nearest settlement" and "modal vegetation height" are most influential in habitat selection at the landscape scale. The inclusion of "total grass cover" and "*Artemisia* cover" in a second model with Δ_i around 2 suggests these factors being crucial for habitat selection as well. Hence, the presence probability of Sociable Lapwing colonies increases with increasing distances to rivers and human settlements, decreases with increasing vegetation height and reaches a maximum at certain coverage of either *Artemisia* or grassy plants (Fig. 9). "Total dung cover" is strongly correlated with "distance to nearest settlement" (r_S=0.71), since both variables indicate the presence of livestock. Habitat selection at the landscape scale is not dependent on topographical factors and soil characteristics according to my data.

Calibration and refinement of the finally achieved averaged model were "very good" according to Backhhaus et al. (2003) with R^2_N=0.915. The model correctly classified 96.4% of all data points. Discriminative power was "outstanding" according to Hosmer and Lemeshow (2000) with AUC=0.996, and "excellent" according to Monserud and Leemans (1992) with Cohen's κ=0.897. Sensitivity and specificity were 0.967 and 0.963 respectively, indicating excellent model performance.

3.3.2 Colony scale

The best model at the colony scale ("nest site selection") achieved by logistic regression with stepwise significance testing identified distance to nearest nest, dung cover, maximum vegetation height and soil cover as the most influent parameters (logit: P(y=1) = -1.505 – 0.007 x dist.nest [+ 0.626 x cov.dung – 0.002 x cov.dung^2] – 0.012 x vegH.max [+0.119 x cov.soil – 0.001 x cov.soil^2]). The model correctly classified 89.5% of all points, and internal validation did not significantly reduce model quality. Overall model performance was good (Nagelkerke's R^2=0.596, $AUC_{bootstrapped}$=0.924, Cohen's κ=0.662).

Using the IT-approach, at the colony scale seven models qualifying as "adequate" were obtained (Appendix A-6), including one model with only one variable.
In the following weighting process, three models remained with w_i different from zero (Table 11). Omitting those with no empirical support (Δ_i>10), at the colony scale two models with acceptable fit remained (Table 11), their weights summing up to 0.998 and thus more than the postulated 0.95.

Table 11: Models with w_i >0, their AIC_C differences to the model with the highest AIC_C (Δ_i) and their empirical support of model i according to BURNHAM and ANDERSON (2002), colony scale.

ID	models	AICc	Δ_i	w_i	level of empirical support of model i
1	cov.dung +cov.dung^2 +cov.veg +cov.veg^2 +dist.nest	215.811	0.000	0.899	substantial
2	cov.dung +cov.dung^2 +cov.veg +cov.veg^2 +vegH.max +dist.nest	220.227	4.416	0.099	considerably less
3	vegH.max +dist.nest	227.949	12.138	0.002	essentially none

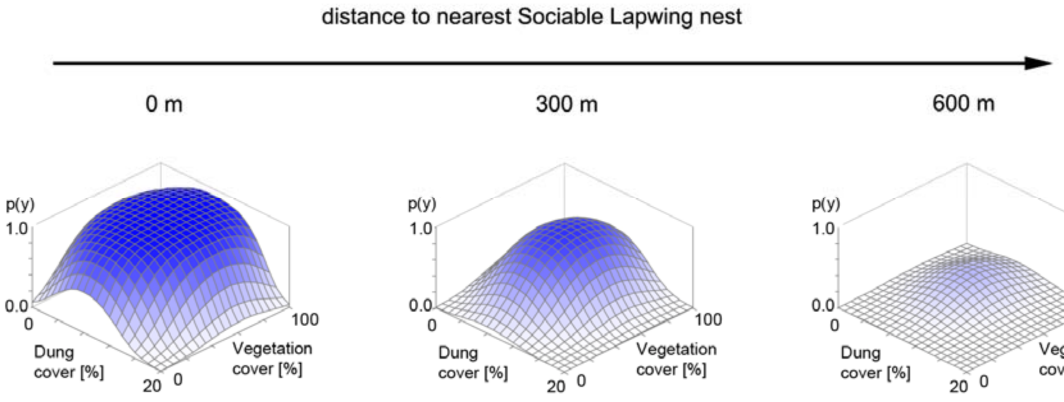


Fig 10. Habitat model with lowest AIC_C ("best model") at colony level. P is plotted against dung cover and total vegetation cover. Diagrams represent different distances from the next Sociable Lapwing nest, blue colour fading out with probability reduction [AUC=0.92, R^2_N=0.56, Cohen's κ=0.70, all values bootstrapped].

Model averaging resulted in three variables with very high weights and one with considerably less weight (Table 12).

Table 12. Averaged coefficients, variable weights and variable ranking for the colony scale model set.

variable	averaged *β*	weight	rank
Intercept	-2.418	1.000	-
dist.nest	-7.358	1.000	1
cov.dung	0.648	0.998	
cov.dung^2	-0.041	0.998	2
cov.veg	0.110	0.998	
cov.veg^2	-0.001	0.998	3
vegH.max	0.056	0.101	4
cov.grass	-0.030	0.000	5

According to both Δ_i values and variable ranking due to AIC weights, is became evident that the presence of other Sociable Lapwing pairs (variable "distance to nearest nest") is the most influential factor in nest site selection, as expected for a colonial species. However, the high rank of the variables "cover of dung", and "total vegetation cover" suggests that these are strongly influencing nest site selection as well. Hence, the probability of nest presence increases with decreasing distance to a neighbouring breeding pair and an increasing cover of dung. It reaches a maximum at a certain cover of vegetation, i.e. Sociable Lapwing prefers areas with a certain amount of bare patches for building nests (Fig. 10). Habitat selection at the colony scale is not dependent on the availability of higher vegetation for rearing chicks according to my data.

Calculated values for discriminative power, calibration and refinement were slightly lower than for the landscape scale models suggesting models being poorer than at the landscape scale.

However, calibration and refinement of the finally achieved averaged model were still “very good” according to BACKHAUS et al. (2003) with R^2_N=0.573. The model correctly classified 86.2 % of all points. Discriminative power was “outstanding” according to HOSMER and LEMESHOW (2000) with AUC=0.919, and still “good” according to MONSERUD and LEEMANS (1992) with Cohen’s κ=0.644. Sensitivity and specificity were 0.867 and 0.861 respectively, indicating good model performance.

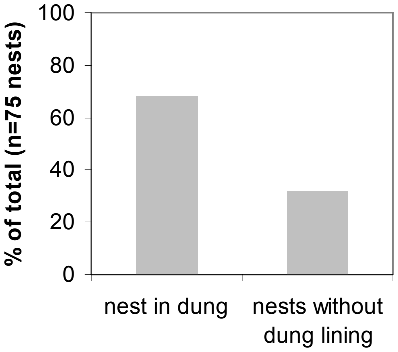


Fig. 11. Proportion of nests found situated in/outside dung piles

Out of all nests considered in variable sampling, 68 % were placed into or situated in direct vicinity of dung piles (Fig. 11, Appendix D-8). This suggests that dung is important in nest site selection.

3.4 Predictive modelling and external model validation

The best multivariate model, i.e. the model with lowest Δ_i, including only variables exhaustively covered for the whole study area at the landscape scale was model 10 (cf. Table 9). This model had “distance to nearest settlement” and “distance to nearest river” as covariates. With AIC_C=49.8 and Δ_i=13.4, it was significantly worse than the best model. However, overall model performance was still very good with R^2_N=0.812, Cohen’s κ=0.895, AUC=0.979 and CCR=0.964 (all values bootstrapped with n=300 iterations).

Applying this model led to a prognosis of suitable habitat distributed patchily across the whole study area (Fig. 12 to 14). It is evident that suitable habitat is available only very localised, and patch distribution scattered. The largest areas of suitable habitat are situated around the settlements Korgalzhyn and Aktubek (cf. 2.1, Fig. 3), both well-known as high-density breeding areas for Sociable Lapwing.

Temporal external model validation with nest sites (n=375) from 2004 (data: W. CRESSWELL pers. comm. 2007), 2005 and 2007 (data: R. SHELDON pers. comm. 2007) led to excellent agreement and conformity: 94.4 % of all nests were situated within areas classified as “suitable” (Fig. 13, 14). However, only 58% of all patches labelled as suitable (i.e. mirroring a occurrence probability >0.5) were occupied in any of the control years. This points to missing covariates in the model (vegetation height not included), but also suggests that there is still “empty habitat” suitable for further colonisation, i.e. carrying capacity is not reached yet in the study area.

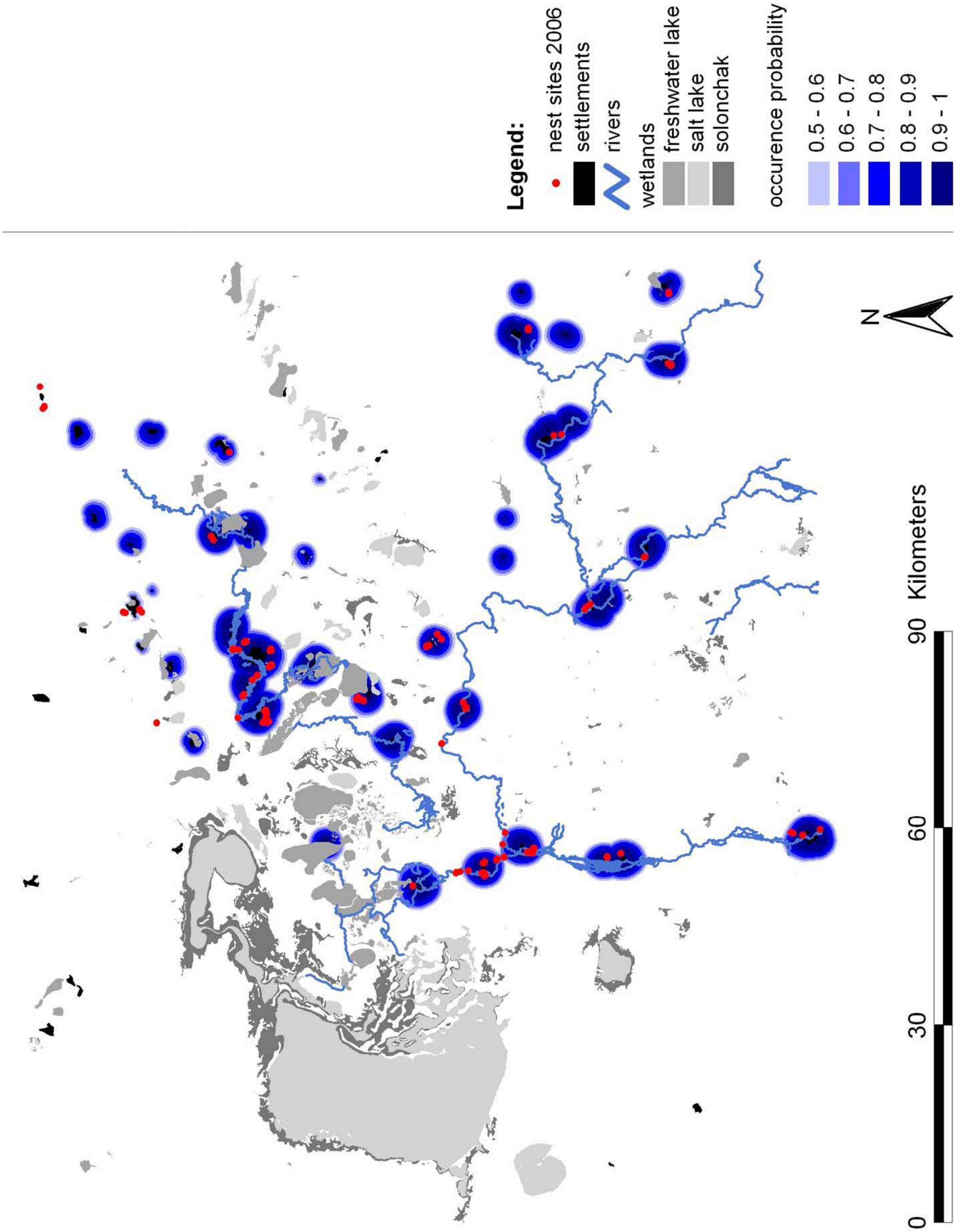


Fig. 12. Habitat suitability map covering the whole study area. Suitable habitat defined as area where occurrence probability is >0.5, based on a model including only “distance to nearest settlement” and “distance to nearest river” as predictors. Dots represent nest sites from 2006 (n= 168).

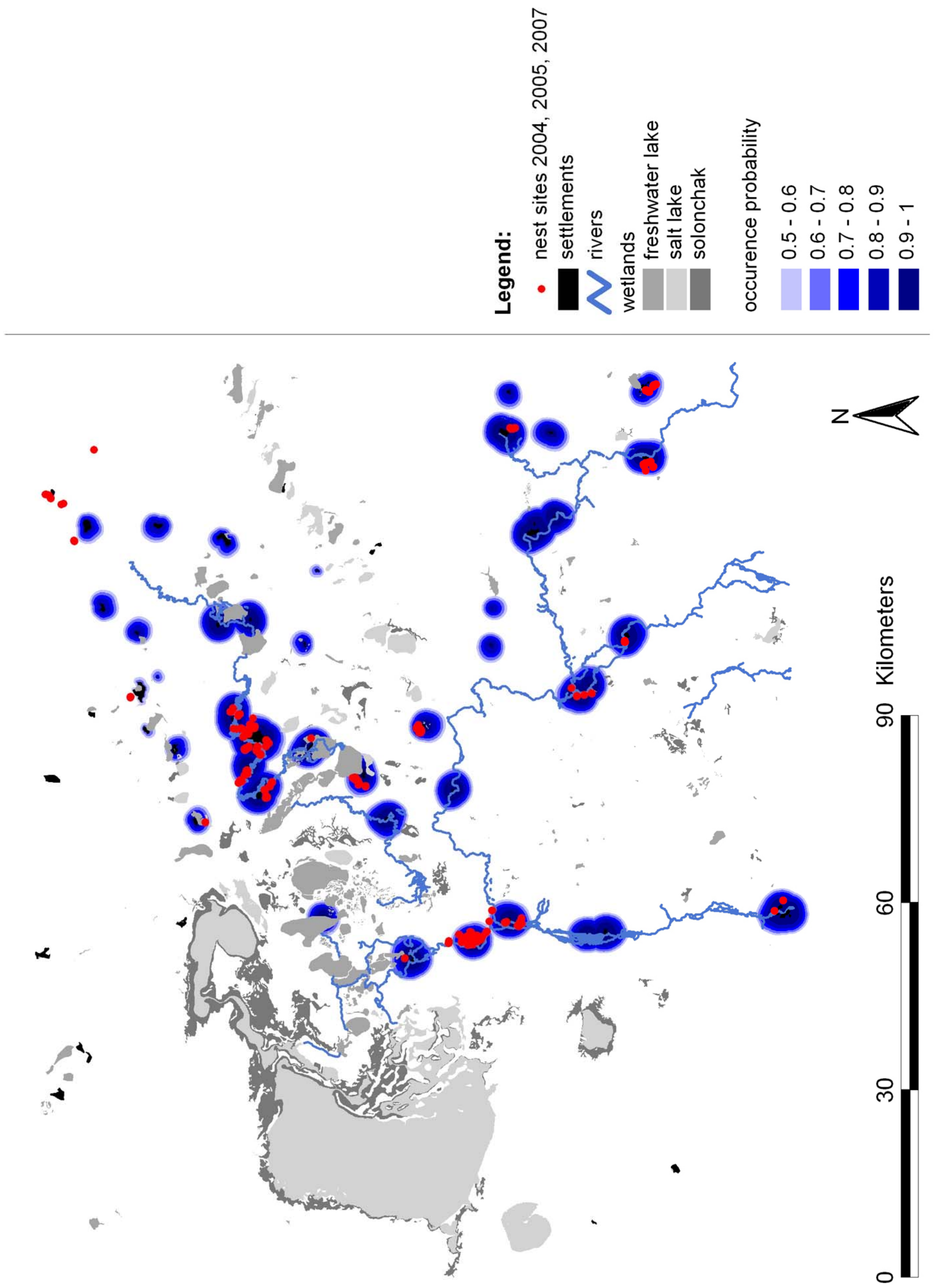


Fig. 13. Habitat suitability map covering the whole study area. Suitable habitat defined as area where occurrence probability is >0.5, based on a model including only “distance to nearest settlement” and “distance to nearest river” as predictors. Test for temporal model transferability: dots represent nest sites from 2004, 2005 and 2007 (n= 375).

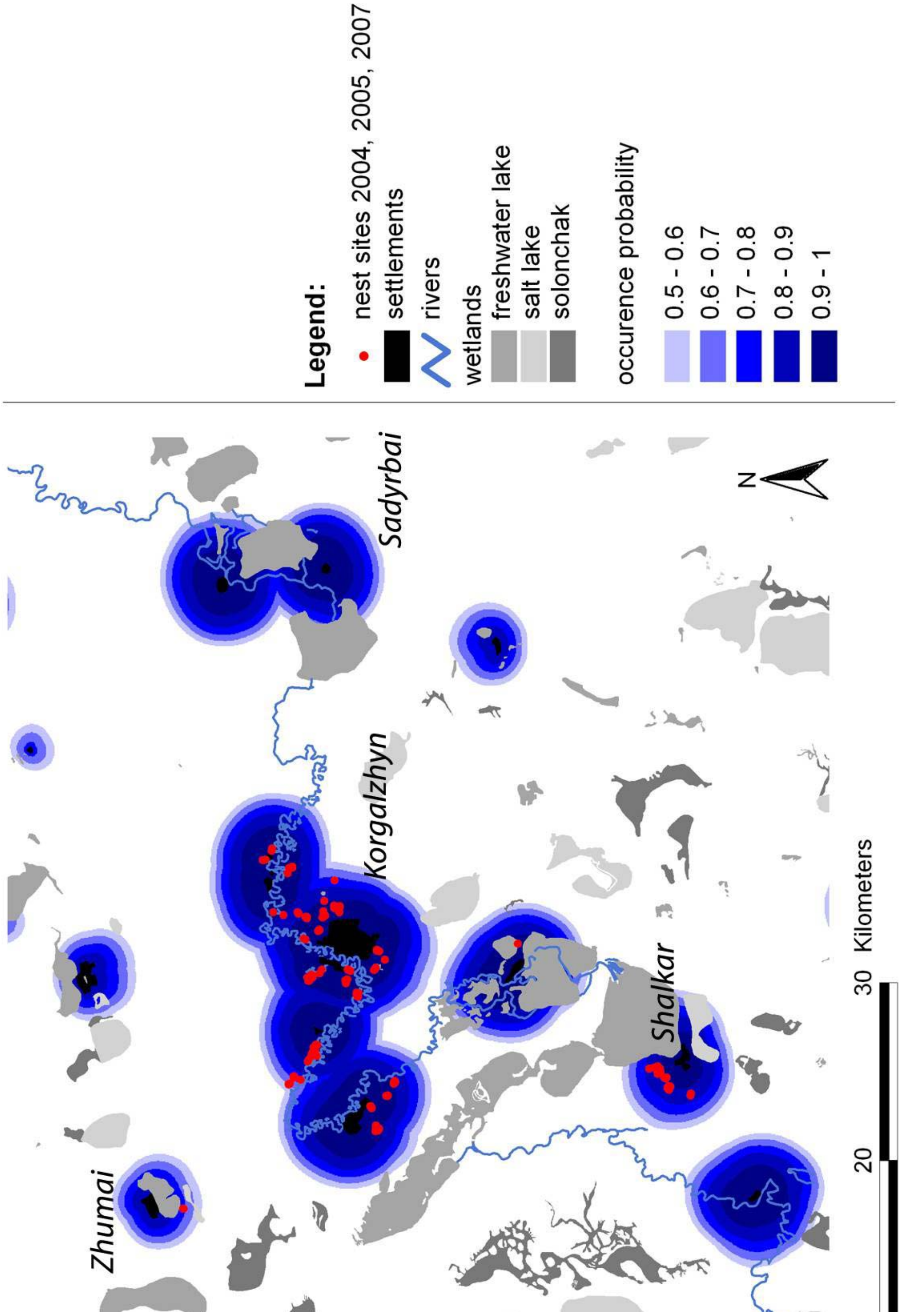


Fig. 14. Detailed habitat suitability map covering main breeding areas of Sociable Lapwing around the settlement of Korgalzhyn. Suitable habitat defined as area where occurrence probability is >0.5, based on a model including only “distance to nearest settlement” and “distance to nearest river” as predictors. Test for temporal model transferability: dots represent nest sites from 2004, 2005 and 2007 (n= 375). Note good temporal model transferability: Most nest sites are situated in areas with occurrence probabilities of 0.8 to 1.0.

4 DISCUSSION

4.1 Discussion of methods

4.1.1 Sampling design and data error

Habitat models are sensitive for two kinds of error: Error resulting from data deficiencies and error introduced by the specification of the model. Regarding the first fact, especially sample size is often an issue of concern (BARRY and ELITH 2006). As shown by HIRZEL and GUISAN (2002), model quality is affected first by sample size, with crucial impact of the sampling strategy as well. The authors achieved acceptable results in predictive modelling with a sample size of n=116 presences/absences, but increased prediction accuracy by 15 % when increasing sample size to n=450 presences/absences, i.e. four times. A further twofold increase in the number of presences/absences did not increase prediction accuracy any more. So, obviously, increased sample size increases the predictive power of a model, but the incremental increase in model performance rapidly decreases with increasing sample size.

Sample size in this study was n=139 presences/absences at landscape level, and n=340 presences/absences at colony level. This seems to be a good compromise between extremely time consuming data collection with more points and unpredictable model behaviour with less. At the landscape scale, time for data collection in some areas reached two hours per point in very rough terrain, thus an increase in sampled absences would have led to an unjustifiable effort. At the colony scale, sample size was large enough to achieve good results according to HIRZEL and GUISAN (2002).

Furthermore, the ratio of presence to absence data is important. Increasing sample size would have been possible only by increasing the number of absences, since I considered all presences available for modelling. This would have led to significant reduction in prevalence, which should not fall below 20 % as suggested by BONN and SCHRÖDER (2001). Model quality would probably have been increased substantially by a higher prevalence, but as colony size was small and colonies scattered, the number of nest sites available for sampling was limited.

Sampling strategies used in point sampling include regular grid sampling with data collection at the grid nodes, completely random sampling, and different random sampling methods with stratification after landscape features (equal-random stratified, proportionally random-stratified) (HIRZEL and GUISAN 2002). Model quality measured as prediction accuracy is theoretically best when applying equal grid or proportional random-stratified sampling (HIRZEL and GUISAN 2002), but there are several examples for good models achieved by completely random sampling (e.g. OPPEL et al. 2004).

I used completely random sampling, as it was not possible to get accurate data on landscape structure prior to fieldwork. There would have been probably no great difference if I had applied random-stratified sampling, as the landscape is quite homogeneous, and most habitat types are

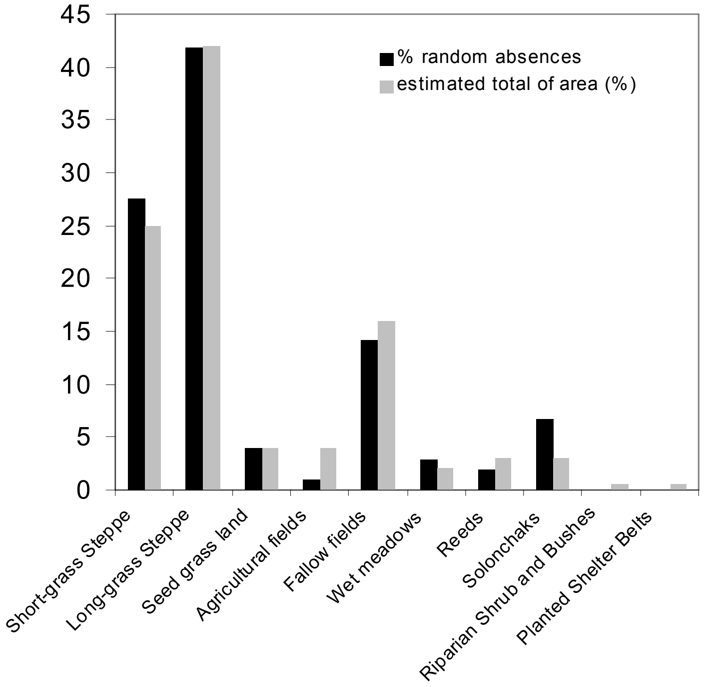


Fig. 15. Proportion of random points falling in each habitat category found in the study area (in %, terrestrial biotopes only, cf. 2.1), and estimated proportion of all habitat types for the entire study area according to Landsat 7 ETM+ imagery. Random points falling in areas of shrub, bushes and planted shelter-belts were not considered during variable recording, thus their proportion here is 0.

found in large shares. At the landscape scale, the habitat type Short-grass Steppe (cf. 2.1 for habitat type definition) has been covered slightly disproportionately, as have solonchaks (Fig. 15). Deviation was most pronounced in used and fallow fields. This was because of some random points were centred on worked fields and not accessible due to the large size of the worked patches (400 ha in average) and land owners' concerns. However, overall deviation was very low, which allows to consider my approach proportionally random-stratified sampling with predefined habitat types as *strata* (Fig. 15). As a conclusion, the sampling strategy was appropriate to achieve good data for modelling.

A further possible error source is the definition of "colonies". It was not possible to demarcate colonies using individual based home-range sizes prior to data gathering. Because of this, a colony was defined as a site with an accumulation (inter-nest spacing less than 500 meters) of at least two breeding pairs at least three kilometres apart from the next breeding incidence (cf. 2.2.3). The spatial extent of each colony was determined by buffering all nest sites cumulatively with r=500 m relying on results of observed individual home-range size from the year 2005 (own data). Colonies were represented by an area of one or several connected circles. However, the area used by all Sociable Lapwings of a colony is certainly not strictly circle-shaped. Thus, in some cases the circles were likely to include habitat not "available" (e.g. JONES 2001) for the breeders and might have added some bias to the data collected.

The choice of habitat parameters to be examined for an influence on Sociable Lapwing habitat selection was made hypothesis-based. I consider the set of covariates included in modelling as the most appropriate choice under the given circumstances. However, there are further factors influencing habitat selection in birds, which are lacking in my approach: Selection of breeding habitat might also be driven by food availability (CODY 1985, MARTIN 1987), as proved for Northern Lapwing (JOHANSSON and BLOMQVIST 1996). Collecting data on food abundance was not possible given the large spatial scale of my study. This might have influenced model quality. However, food abundance is probably correlated with the abundance of large grazers, and thus covered by the selected variables “distance to nearest river” and “dung abundance”. This is a typical case of modelling with indirect or “substitute” variables (BARRY and ELITH 2006), but I cannot judge to what extent information is lost when not considering the proximate factors such as arthropod densities.

4.1.2 Accuracy of bird data and parameter recording

I aimed to get every Sociable Lapwing clutch in the study area as a presence point during the considered first clutch incubation period from 23rd April to 25th May to facilitate an analysis of comparable samples (entirety of nest sites vs. entirety of random absences). Surveys for colonies and clutches were conducted by three project field teams with two field workers each during the entire period, fairly well organized in terms of logistics (communication, transport). Given an average daily work period of 10.5 hours, total survey effort was 1,040 man-hours during the mentioned time period. Following these circumstances, survey effort was maximised. However, there is evidence that some breeding pairs were missed, as in June females with chicks were observed in regions were no clutches had been found before. From observations in all colonies during the chick-period, I estimate that the project teams found approximately 95% of all clutches produced within the given period. As this is equal to the lower limit of a 95% confidence interval of estimated clutch numbers, I suggest the overlooking of some clutches to be a minor source of sample incomparability.

As there were three teams of fieldworkers surveying different parts of the study region, one would assume observer-induced bias in regional survey coverage. Since survey teams were not restricted to certain areas, but every team worked at every point of the study area during the survey period, observer-induced bias is probably of minor importance.

There is evidence that I missed some Rook colonies situated in remote willow thickets along rivers or consisting of less than ten pairs. Judging from habitat availability, I estimate that I detected approximately 70 - 80 % of all colonies within the study area, and counted approx. 90 % of all Rook breeding pairs. Missing some colonies might influence the values of the variable “distance to nearest rookery” for both absence and presence points. I did all rookery surveys myself and checked those discovered by others again, thus there is no observer-induced bias to be expected.

Biased samples can affect both variable selection and coefficient estimation in regression models (MACKENZIE et al. 2002, TYRE et al. 2003). Thus, it needs to be checked, if systematic error in variable recording might have occurred.

All topographic variables were recorded with standard equipment and GPS units. Under the given topographical and weather conditions (low cloud cover) GPS receivers proved to be precise to 4 meters in horizontal and to 10 m in vertical direction (own tests, n=150 readings). Maximum outliers were 15 m deviation in horizontal, and 45 m deviation in vertical direction. This constitutes a suitable resolution for topographical variable recording in the case of my study.

Plant composition and coverage was recorded at 2 x 2 m sample plots, centred on the nest sites and absence points. Coverage was estimated as percentage precise to the nearest 5 %, the level of estimate accuracy was checked by repeated sampling prior to the fieldwork period. As deviation in repeated samples was always below 10 % independent of observer, I regard this method as adequate. It is unlikely that the models suffered from systematic cover estimate errors. As I conducted more than 50 % of the samples on my own and controlled every sampling myself, there is no observer-specific bias to be expected. However, it is not clear if the use of 4 m² plots is appropriate for vegetation characterisation, or if a sound approach requires larger plots. Botanists working in the region use 20 x 20 m plots for phyto-sociological cover estimates (T. DIETERICH, F. SCHWAHN pers. comm., 2006) due to the patchy character of steppe associations in some areas. It seems possible that the features measured at a presence/absence point represent a situation not typical for the surrounding landscape patch, although from field experience I restrict these cases to less than five percent of the total sample.

Since vegetation growth within the recording period would limit comparability of single presence/absence points and model quality, I tested whether this was an issue over a period of 15 days. I repeated vegetation height measurements at randomly chosen points out of the predetermined presences and absences once (first date 21st May, repeated measurement 05th June 2006). Neither modal nor maximum vegetation height differed significantly (Wilcoxon-$Z_{\text{modal height}}$= -1.358, $p_{\text{modal height}}$=0.175; Wilcoxon-$Z_{\text{maximum height}}$=-1.068, $p_{\text{maximum height}}$=0.285, n=63, Fig. 16).

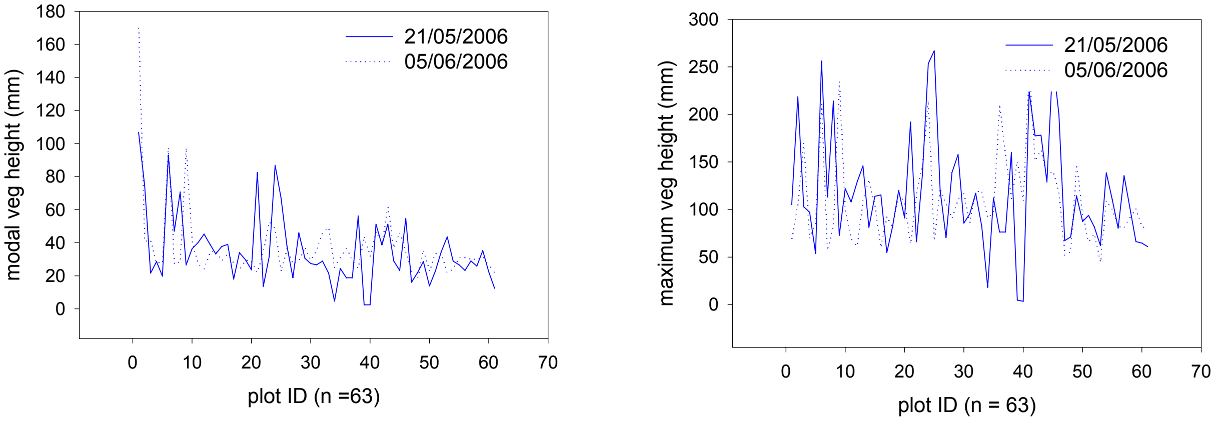


Fig. 16. Modal (left) and maximum (right) vegetation height during first and repeated measurements at 63 absence and presence points.

4.1.3 Modelling procedure

There are several methods available for the analysis of spatial data with regard to habitat selection. In the case of Sociable Lapwing, three types of statistical data analysis allowing predictive modelling were considered suitable beforehand: Multiple Discriminant Analysis (MDA, cf. FISHER 1936, FIELDING and HAWORTH 1995), Multiple Logistic Regression (LR, overview in HOSMER and

LEMESHOW 2000) and Artificial Neuronal Networks (ANN, e.g. ÖZESMI and ÖZESMI 1999). Comparing all three in a case study with river birds, MANEL et al. (1999) found logistic regression to outperform both remaining approaches in both model performance and prediction accuracy. SCHRÖDER and REINEKING (2004) analysed the current popularity of several modelling approached and found an increasing predominance of logistic regression approaches along with a significant decrease of discriminant analysis for the last ten years. Neuronal Networks were increasingly used as well, but on a very low level of publications.

Relying on these hints from literature, I chose logistic regression as modelling strategy. This had further advantages, as this method proved to be relatively robust concerning deviation from normal distribution (BACKHAUS et al. 2003) and allows the inclusion of a few categorical variables. It gives much better results compared to presence only models (BROTONS et al. 2004, ENGLER et al. 2004): Models lacking absences often inaccurately identify attributes of unsuitable sites, hence presence-absence modelling is suggested by most authors (cf. BARRY and ELITH 2006, BROTONS et al. 2004). Models considering species' abundance in many cases provide little additional information when trying to identify parameters influencing habitat selection (PEARCE and FERRIER 2001). Exceptions are models considering species density in relation to habitat parameters and connectivity in fragmented landscapes (e.g. MICHALEK et al. 2001 for woodpeckes, a group with very low mobility). Furthermore, in the case of my study it would have been difficult to gather appropriate abundance data, since Sociable Lapwing colony size was overall low (largest colony held 12 pairs) and did not differ noticeably across the colony sites.

4.1.4 Stepwise significance testing and IT approach – a comparison

I used two strategies of model simplification: a stepwise backward approach, and an information theoretic approach based on the comparison of all models possible to build. For comparison of these two methods cf. 2.3.1.

Both approaches gave qualitatively similar approaches. At the landscape scale ("breeding habitat selection"), the best model achieved by logistic regression with stepwise significance testing identified "distance to nearest settlement", "distance to river" and vegetation height as the most influent parameters. The same factors were also suggested by the model with the lowest Δ_i ("best model") achieved with the alternative procedure. At the colony scale ("nest site selection"), the best model achieved by logistic regression with stepwise significance testing identified "distance to nearest Sociable Lapwing nest", dung cover, vegetation height and cover of bare soil as the most influent parameters. The model with the lowest Δ_i ("best model") achieved with the alternative procedure was however more parsimonious and included only "distance to nearest Sociable Lapwing nest", dung cover and cover of bare soil (as variable "vegetation cover"). A model including the additional variable vegetation height did not much worse (AIC_C decreased by 4.4, here indicating a less parsimonious model). To sum up, in the case of Sociable Lapwing there was not much difference in results gained by stepwise significance testing vs. AIC comparison. Large divergence is to be expected only, when there is little difference between the impact of single habitat variables, or if parameter selection does not match key mechanisms, but only marginally important features. As the IT approach allows parameter ranking, the hierarchical structure in habitat selection can be better estimated (cf. 4.2.2).

4.1.5 Model generality and transferability

The models achieved in this study do not only quantify the influence of different factors in Sociable Lapwing habitat selection, but also allow spatial predictive modelling. With the information available, habitat suitability maps can be created for all areas where information about the included parameters is available spatially exhaustive – assuming similar habitat preferences in all populations.

FIELDING and HAWORTH (1995) showed that spatial model transferability might be very low, when interregional differences in nesting habitat are high. Modelling habitat selection of three large avian predators, namely Golden Eagle (*Aquila chrysaetos*), Raven (*Corvus corax*) and Common Buzzard (*Buteo buteo*), they found the interval in correct prediction of areas suitable for nesting ranging form 8% to 100%, dependent on data recording scheme and regional focus.

Although I achieved a habitat suitability map for the study area for only two variables, namely “distance to nearest settlement” and “distance to nearest river”, overall model performance was very good with $R^2_N = 0.812$, Cohen’s $\kappa = 0.895$, AUC = 0.979 and CCR = 0.964 (all values bootstrapped with 300 iterations) indicating a suitable spatial prediction accuracy. However, there were many areas predicted as suitable, where there were no colonies of Sociable Lapwing found from 2004-2006, evidently mainly because of unsuitable vegetation structure. Thus, the prediction was still quite coarse due to the lack of an important variable.

In general, one might conjecture the achieved models as quite suitable for prediction, as they identified factors similar to those mentioned in earlier, more anecdote studies. However, predictions using the models achieved in this study should be done only in areas with similar landscape structure as in my study area. I cannot judge on overall spatial transferability of the models – this has to be tested during a subsequent fieldwork season.

Transferability in time, i.e. temporal generality was tested by applying the model used for predictive modelling to nest sites found in 2005. 98.8 % of all nest sites from that year were situated in areas predicted as “suitable”, so transferability in time was good. This means that habitat features change little, and spatial dynamics of suitable habitat patches are low between years. Therefore, the model can be considered suitable to be applied in subsequent years as well.

4.2 Discussion of results

4.2.1 Habitat use

Main habitats used by Sociable Lapwings in the study area comprised Short-grass Steppe, fallow wheat and seed grass communities usually dominated by Wormwood *Artemisia*, recently (autumn 2005 and spring 2006) burnt Long-grass Steppe (predominantly Feather Grass *Stipa*-communities) and dry Solonchaks with dense *Salicornia* cover (Fig. 6).

The preference for short-grass steppe and semi-desert is common sense in most of the references I reviewed (overviews in CRAMP and SIMMONS 1983, DEMENT'EV and GLADKOV 1951, DOLGUSHIN 1962). We found most colonies (with 52% of all nests found) in strongly-grazed *Artemisia*-, or *Artemisia-Festuca*-dominated Short-grass Steppe, often close to settlements. This was previously mentioned by a row of authors, especially SAMIGULLIN (1987), SOLOMATIN (1997), BRAGIN (1999), GORDIENKO (1991) and KHROKOV (1977, 1996). SOLOMATIN (1997) and BEREZOVIKOV et al. (1998) gave special emphasis to the fact that colonies were often situated close to human settlements.

The second-large proportion of “near-natural” (never ploughed up) habitat (with 14% of all nests found) concerns semi-desert type Short-grass Steppe, moderately grazed, but due to climatic conditions sparsely vegetated and clearly dominated by *Artemisetum*-communities. These areas are typical within the transition of steppe to semi-desert, and Sociable Lapwing regionally is still abundant within this habitat type (DAVYGORA et al. 1989, GORDIENKO 1991, BEREZOVIKOV et al. 1998, RYABOV 1974, pers. obs. from 2006 surveys in Dzheskazgan region, Southern Central Kazakhstan).

On agricultural land, most nest sites (26% of all nests found) within the study area were situated on fallow seed grass (*Agropyron cristatus*), and only 2% of all nests were found on fallow cereal fields. Both fallow seed grass and cereal fields serving as breeding habitat for Sociable Lapwing were mentioned in a variety of publications (e.g. BELIK 1998, 1999, BRAGIN 1999, KHROKOV and KARPOV 1999, KARYAKIN and KOSLOV 1999). Despite intensive search on recently ploughed or sown-in fields in the beginning of May, there was no evidence for Sociable Lapwing breeding in these areas. A group of four males and five females displaying on a recently ploughed, vegetation-less field in mid May moved 2 km south when cereal growth started and bred on strongly grazed Short-grass Steppe (pers. obs.). In Russia, where there is most of the steppe ploughed up now, breeding of Sociable Lapwing on ploughed or sown-in fields is regularly observed (BELIK 1998, BRAGIN 1999, KARYAKIN and KOSLOV 1999, IL'ICHEV and FOMIN 1979, SOLOMATIN 1997, MOROZOV 2005, with photograph).

Solonchaks covered with sparse vegetation (usually dominated by *Salicornia* spp.) were hardly used within the study area with only 4% of all nests found. Recent publications on Sociable Lapwing from the Korgalzhyn study area did not mention breeding incidences on Solonchaks (EICHHORN and HEINICKE 2000, EICHHORN and KHROKOV 2002, WATSON et al. 2006), but there are several hints from other papers (GRACHEV 1973, GORDIENKO 1991, KUCHIN and CHEKCHEEV 1987, SHEVCHENKO et al. 1993), although the features of the described “Solonchaks” stay rather unfocused.

In spite of exhaustive literature search I did not find any published case of Sociable Lapwing breeding on burnt steppe, although numerous observations indicate that the birds use this habitat during other seasons (overview in DEMENT’EV and GLADKOV 1951). We found birds breeding on recently (April 2006) burnt *Stipa*-dominated Steppe with very low (<5%) *Artemisia* cover. Other colonies in this habitat were detected by members of the Sociable Lapwing research and conservation project in Pavlodar region, NE Kazakhstan in spring 2006 (V.V. KHROKOV, A. KNIGHT in litt., 2006).

Other habitat types described before as used by Sociable Lapwing for breeding include:

- “pristine steppe” and *Stipa-Festuca*-Long-grass Steppe (KHROKOV 1977, SAMIGULLIN 1987, VOLCHANETSKII 1937, RYABOV 1974)
- mountain steppes (Southern Ural area, DAVYGORA et al. 1989)
- *Haloxylon*-(Saxaul-) Semi-desert (GRACHEV 1973)
- Sandy soils and hilly sand-dunes with Quack grass (*Elymus giganteus*, *E. repens*) (VOROB’EV 1936, RYABOV 1974)

I did not find any breeding colonies in “pristine steppe” or the probably equivalent *Stipa-Festuca*-Long-grass Steppe, but rather detected a clear avoidance of this habitat (c.f. 4.2). The remaining habitat types are missing within the study area.

If near-natural breeding habitat is not available, Sociable Lapwing occupies a variety of secondary habitats. Breeding was confirmed close to an airport runway, a vegetation-less building site at the outskirts of a big city, and at short, perished vegetation close to the dump of a chemical plant (SOLOMATIN 1997). In my study area, these “industrial biotopes” are not to be found, and breeding is restricted to more or less natural areas and agriculturally used sites.

4.2.2 The hierarchical process of habitat selection in Sociable Lapwing

As habitat selection is a hierarchical process of behavioural responses, neither proximate nor ultimate factors can be determined by simply studying current bird distributions (CODY 1985, JONES 2001). The results presented can provide only hints on this process. Variable constellation of the best models and the ranking of habitat variables suggest that breeding habitat selection at **landscape level** (hierarchical level 1) might be directed as follows: Increased occurrence probability close to rivers (variable “distance to river”, weighing rank 1, cf. Table 10) might by influenced by the fact the birds migrate mainly (but not only!) along rivers (orientation strips and water supply in the monotonous steppe landscape). Migration along rivers is proved by data from colour-ringed and satellite-tagged birds from Kazakhstan (R. SHELDON in litt. 2007, pers. obs.). As birds do not migrate evenly across the landscape, suitable habitat along rivers is probably easier discovered and occupied with a higher probability than areas far off rivers.

When the birds arrive at sites with short vegetation (variable “maximum vegetation height” has maximum value, rank 2), they go down to inspect the area, but might stay only if dung density is sufficient (variable “total dung abundance”, rank 6, is highly correlated to and represented by “distance to settlement”, rank 3). High dung densities undoubtedly increase invertebrate abundance (TUCKER 1992, ATKINSON et al. 2004), so stronger grazing usually means higher food abundance. The driving proximate factors are not clear: either birds “rely” on their experience that short swards mean more dung and thus a better food basis, or they visually inspect the areas looking for dung piles. The latter is suggested by the high proportion of clutches placed into dung (cf. 3.3.2)

Occurrence probability is higher, if there is water close by, indicated by the variable “distance to nearest river” and “distance to water” (rank 4). As “distance to nearest river” and “distance to nearest settlement” are not significantly correlated (r_s = -0.132), higher occurrence probability closer to rivers is not an artefact resulting from settlements being situated mostly close to rivers. In contrast, the preference for areas with a high cover of *Artemisia* spp. (variable on rank 5) as well as the avoidance of areas with high grass cover (rank 7) are probably artefacts: *Artemisia* is selected positively when grazing pressure is high (as it is relatively unpalatable), and grasses decrease in coverage in this case. This is also suggested by the fact that both variables are correlated (negatively resp. positively) with vegetation height.

Within the study area, habitat selection at the landscape scale is thus mainly directed by three factors: the distance to water and especially rivers, the presence of livestock and a suitable vegetation height. This is obvious, as the coefficients of the variables indicating these features have a considerable higher weight than all others in model averaging, and are exclusively selected in stepwise regression models.

Nest site selection at **colony level** (hierarchical level 2) might subsequently occur as follows: First birds arrive in the areas selected on a higher level, choose a breeding spot in an area with suitable vegetation height, a suitable proportion of bare ground/vegetation and suitable dung density (variables “maximum vegetation height” rank 3, “cover of vegetation”, “dung cover”, both on rank 2). They are possibly directed by the experience that hatching success is highest in these areas (WATSON et al. 2006), but might simply avoid areas with tall and dense vegetation due to the experience that there prey detectability is lower. As nests were situated in or near dung piles in two thirds of all cases (see below), one might conjecture that not prey abundance, but immediate dung pile number influences the selection of breeding sites. The males start scratching out scrapes at these areas, then mated females (often arriving with inseminated eggs in the oviduct, DOLGUSHIN 1962) choose one of these spots and start egg laying (e.g. SOLOMATIN 1997). Later arriving birds might not primarily look for suitable habitat features, but use public information for nest site selection. Public information is very important in a row of species (e.g. DANCHIN et al. 1998), and it has recently been proved that some species decide where to nest and whether to return the next year based in part on knowledge of their neighbours' reproductive success (DOLIGEZ et al. 2002).

This would explain small scale distribution patterns. Hence, occurrence probability decreases with increasing distance to the next nest (variable "distance to nearest neighbouring SL nest", rank 1).

The only study available on habitat selection of Sociable Lapwing examined nest site selection at the colony scale in Korgalzhyn region (WATSON et al. 2006), but results are based on low sample size. The authors found that nests were sited significantly more often in areas dominated by *Artemisia* spp. rather than grass and with higher percentage of dung cover than points 500 m apart from the nest. Vegetation height and cover of dung were not significantly different at nest sites and 500 m away from nest. In spite of the consideration of only four variables, the results point in a very similar direction as my data.

As 68 % of all nests were placed into or situated in direct vicinity of dung piles (cf. Appendix D-8), there is direct evidence that dung is important in nest site selection. This might be caused by higher food abundance (scatophagous insects and *Diptera* attracted to grazers), as birds have been seen pecking insects from fresh and decaying dung piles (pers. obs., cf. also TUCKER 1992, ATKINSON et al. 2004). Other hypotheses include a possible heating effect of decaying dung, describe dung piles as orientation marks to facilitate recovering of the own clutch in a colony environment, and, most likely, suggest dung piles to have a camouflage effect on embedded clutches.

I did not find a relationship between slope, inclination and height above the sea level on the one hand, and Sociable Lapwing occurrence probability on the other. This strongly suggests that positive selection of southern exposed slopes or areas with distinct micro-relief (e.g. SOLOMATIN 1997) is not true. Also, I could not confirm claims on preference of sandy soils (KHROKOV 1996) and avoidance of broken terrain and rough, stony surface (CRAMP & SIMMONS 1983).

4.2.3 Is Sociable Lapwing population development mainly linked to habitat availability?

Undoubtedly, Sociable Lapwing numbers strongly decreased until the 1990s (EICHHORN and KHROKOV 2002, cf. Introduction). There are, however, several hints that numbers of Sociable Lapwing have been increasing again since at least the year 2000. SHELDON et al. (2006) conducted rapid surveys over huge areas in Kazakhstan during the breeding season and extrapolated a new estimate of 5,600 pairs worldwide, being substantially higher than all previous estimates since the end of the 1980s (cf. 1.1). Large flocks migrating on the Middle East flyway (altogether more than 2,600 birds synchronously counted in early spring and 3,200 birds only in Turkey in autumn 2007, R. Sheldon pers. comm., BIRDLIFE INTERNATIONAL 2007) as well as good breeding success in Central Kazakhstan for the last two years (SHELDON et al. 2005, R. SHELDON pers. comm., 2006) point to the fact that the severe decline reaching its maximum at the end of the 20th century halted. What are the reasons for this development?

Changes in land use and livestock numbers My research revealed that, at least in Central Kazakhstan, Sociable Lapwing occurrence is strongly linked with land use and especially livestock breeding today. Only a minority of birds were found to nest in near-natural habitat (solonchaks, burnt areas), and features preferred in habitat selection according to this study (low vegetation, livestock) are found mainly in anthropogenic habitat.

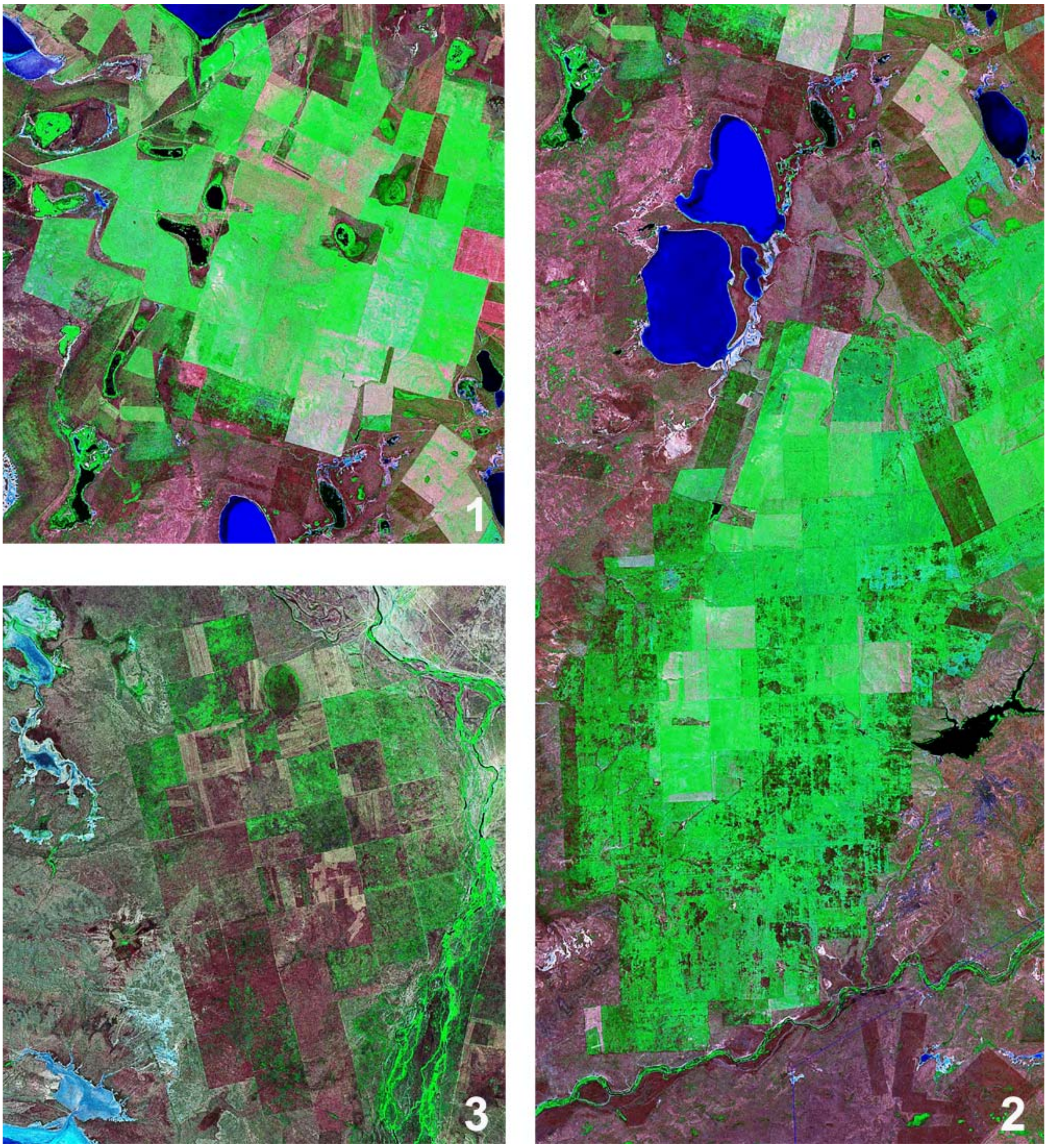
For a livestock-dependent species, the amount of suitable breeding area must be influenced immediately by socio-economic factors as livestock density. A thorough conservation approach linking population development with habitat preferences has to regard these socio-economic features as well.

Knowing the factors influencing habitat selection, and considering the results on habitat use of this study, I suggest that Sociable Lapwing has profited from changes in agriculture, land use practices and landscape change since the major depression in numbers at the end of the 1990s. After the break-up of the Soviet Union, huge areas of formerly cultivated cereal fields fell fallow due to a cut of subsidiaries going along with the end of the state farm ("Sovkhoz") system. DE BEURS and HENEBRY (2004) analysed Normalized Difference Vegetation Index (NDVI) changes by time series statistics and remote sensing across several study areas in Northern Kazakhstan. They revealed an enormous ncrease in the area of fallow land across the whole northern part concentrated on Akmolinskaya oblast', where the study area of my research is situated. Large scale changes in agricultural practise become also evident when analysing total wheat harvest in Kazakhstan, which declined from around 16 Mio. tons in 1988 to about 6 Mio. tons in 1999, equalling a decrease of 62.5 % (MENG et al. 2000).

Livestock numbers crashed after the break-up of the Soviet Union. In Russia, total livestock numbers (cattle, horses, sheep and goats) fell from 57 Mio. head in 1991 to 28 Mio. head in 2000, equalling a decline of roughly 50% (GOSKOMSTAT 2001). In Kazakhstan, the political and economical changes affected livestock numbers as well, resulting in a decline of 58% (10 Mio. to 4.2 Mio head) in cattle and 75% (37 Mio. to 9 Mio. head) in sheep and goats from 1990 to 1999 (SULEIMENOV and ORAM 2000, ROBINSON and MILLNER-GULLAND 2003). This figures are fairly representative for my study area (cf. maps in LENK 2001). A decline in Sociable Lapwing numbers just in this time period (EICHHORN and HEINICKE 2000, EICHHORN and KHROKOV 2002) might well have been linked to the development in land use and livestock numbers.

Since 1999, livestock numbers are increasing again. People in Korgalzhyn region are slowly extending their herds, and there can be observed large flocks of up to 1,000 horses and 800 cattle across pastures in whole Central Kazakhstan again (pers. obs., 2006).

In Korgalzhyn region, 60 % of all formerly agriculturally used area fell fallow between 1989 and 1998 (own analysis from Landsat 7 ETM+ pictures, cf. Fig. 17, p. 47). Sociable Lapwing has obviously profited from this development, since some large areas formerly used for seed grass cultivation are now heavily grazed by the increasing livestock flocks converting them into short, wormwood (*Artemisia*-) dominated swards (Fig. 18, p. 48). Four larger colonies (Amangel'di, Shalkar, Ushsart and Izendi) in the study area are situated entirely or almost entirely in former seed grass cultivation areas, and there was presumably no suitable breeding habitat for Sociable Lapwing around before approximately 1998. In total, 28% of all nests were found in areas, where seed grass and/or wheat has been cultivated until the mid 1990s.



🟩 Cereal field. Recently sown in barley or wheat

🟩 Young fallow field. Fallow for 0-5 years, early successional stage

🟪 Old fallow field. Fallow for 5-15 years, late successional stage in transition to near-natural *Stipa*- steppe

Fig. 17: Current state of three larger wheat cultivation areas within the study area, year 2004:

1) Arykty Sovkhoz, Akmolinskaya oblast' – nearly entire area sown in recently with wheat
2) Korgalzhyn Sovkhoz, Akmolinskaya oblast' – approximately 50% fallow, especially in the southern part
3) Kulanotpes Sovkhoz, Karagandinskaya oblast' – entire area fallow, many old fallow fields (violet) converting into near-natural steppe.

In 2006, even more areas were abandoned compared to these shots.

Landsat 7 ETM+ picture (2004), scale: 1:250,000, north-referenced.

As a conclusion, the overall timing of the mentioned processes both across Northern Kazakhstan and within the borders of my study area supports a hypothesis of increased availability of suitable habitat, and thus a possible halt of the species' decline or even an increase in numbers since the year 2000.

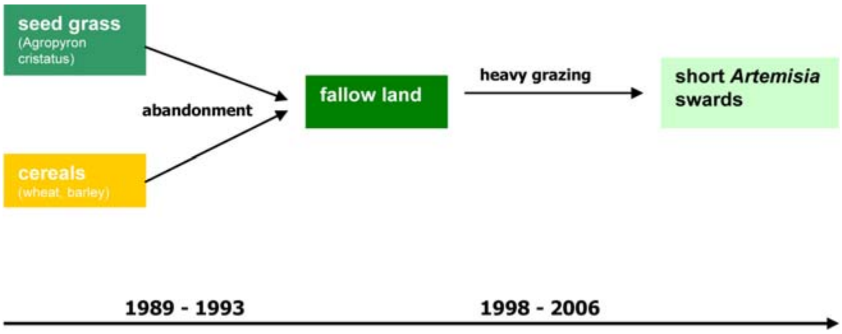


Fig. 18: Changes at agricultural land due to changing land use and agricultural abandonment in Korgalzhyn region. After own analysis of Sovkhoz (State Farm) maps (ZELINGIPROZEM 1985-1988), topographic maps 1:100,000 and Landsat 7 ETM+ satellite images.

Increase in steppe fire numbers Fire is a natural part of the steppe ecosystem. In Kazakhstan, 1,543 fires were observed burning 216,950 ha of forest and steppe in 1997 (KHAIDAROV and ARCHIPOV 2001). Most fires are human-induced, but shares for large steppe fires due to natural reasons reach 35% in some regions. However, in Akmolinskaya and Karaganda oblasts, where the study area is situated, only 0.5% resp. 0.7% had natural origin, especially lightning (KHAIDAROV and ARCHIPOV 2001). A strong increase was observed in area burnt by steppe fires in Kazakhstan from 1981 to 2000 (Fig. 19, after data from KHAIDAROV and ARCHIPOV 2001), mainly due to the worsened economic situation and consequent stop of fire fighting by state fire guards after 1993 due to the fraught economic situation. According to local fire brigades, this trend was pronounced in Korgalzhyn region as well and the area burnt by steppe fires has increased after the year 2000 across the study area. Own observations confirmed this, and after own GIS area analysis, in 2005 approximately 0.8% of the whole study area burnt down (cf. Appendix D-5).

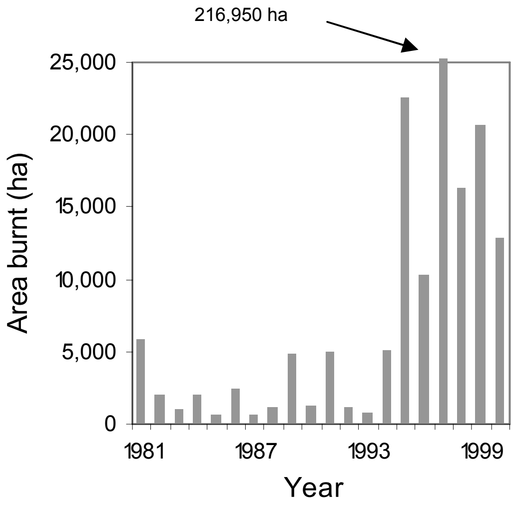


Fig. 19. Area burnt by steppe fires in Kazakhstan between 1981 and 1999 (after data form KHAIDAROV and ARCHIPOV 2001)

As fresh vegetation growth on burnt areas leads to suitable conditions for Sociable Lapwing breeding (low, sparse vegetation, high percentage of open soil), these sites might be considered as primary habitat for the species. Furthermore, freshly burnt areas are often used for livestock grazing, since the fresh *Stipa-Festuca* regeneration vegetation is nutrient-rich (REDMANN et al. 1993). Perhaps these areas attract wild ungulates as well. Both would support an increase in Sociable Lapwing numbers. However, burnt steppe is far less used by Sociable Lapwing as short grazed swards (cf. 3.1), hence the importance of the increase in steppe fires should not be overestimated.

Population development of Saiga antelope (Saiga tatarica) and Sociable Lapwing WATSON et al. (2006) suggested a correlation between the enormous decline in numbers of wild ungulates, especially Saiga antelope *Saiga tatarica* across the Eurasian steppe zone (cf. 1.4), and Sociable Lapwing population decrease. The authors conjectured that Sociable Lapwing evolved to nest in the short swards left in the wake of enormous herds of Saiga and switched to herds of domestic ungulates later on. However, it is doubtful if Saiga herds were ever capable of creating large areas with very short swards as preferred by Sociable Lapwing (optimum below 50 mm as revealed by my analysis). Within the summer distribution area (which largely fits to the distribution range of Sociable Lapwing, SOKOLOV and ZHIRNOV 1998, BEKENOV et al. 1998), the antelope herds consume only 12-30 kg of vegetation per hectare (1.5-5.3% of the pastures yield), because it constantly moves during grazing. The herds do not adversely affect the productivity of their pasture, i.e. do not change species composition and vegetation height on a long time scale. Even at calving places, the animals are highly mobile and do not stay for a longer period in the same area (BEKENOV et al. 1998).

Furthermore, Saiga reach the steppe zone in most years in mid June, in some years already in mid May. This is much too late to create suitable habitat for Sociable Lapwing, which starts incubation often as early as in the end of April (SHELDON et al. 2005).

As a conclusion, I suggest to consider mainly sparsely vegetated solonchaks and recently burnt areas due to steppe fires as **primary habitats** of Sociable Lapwing. The latter are especially suitable for breeding when they were used by wild ungulates such as Saiga antelope for grazing, but the presence of Saiga antelope herds alone was probably not sufficient to create enough suitable habitat for Sociable Lapwing. Nowadays, primary habitats are largely vacated due to the more convenient conditions at pastures of domestic livestock, and Sociable Lapwing turned to a human-following species.

5 CONCLUSIONS and PERSPECTIVES

5.1 From habitat selection to conservation

Although Sociable Lapwing seems to be over the hump and probably does not face immediate danger of extinction, the very small population (cf. 4.2.3) seems still very vulnerable – there is no doubt that the species suffered from an enormous decline probably since the 1890s. It is still unclear, which are the main threats to Sociable Lapwing. It seems, as if low breeding success to what reasons ever was not a crucial problem during the last two years (SHELDON et al. 2005, R. SHELDON pers. comm. 2006), although WATSON et al. (2006) found low breeding success to be pronounced and related that to massive livestock trampling of clutches at breeding areas close to settlements. This would mean, that changes in livestock breeding – from a nomadic way of herding to a more sedentary with livestock being constantly present at the colonies – contributed to decreased breeding success due to increased clutch loss (for details, cf. WATSON et al. 2006).

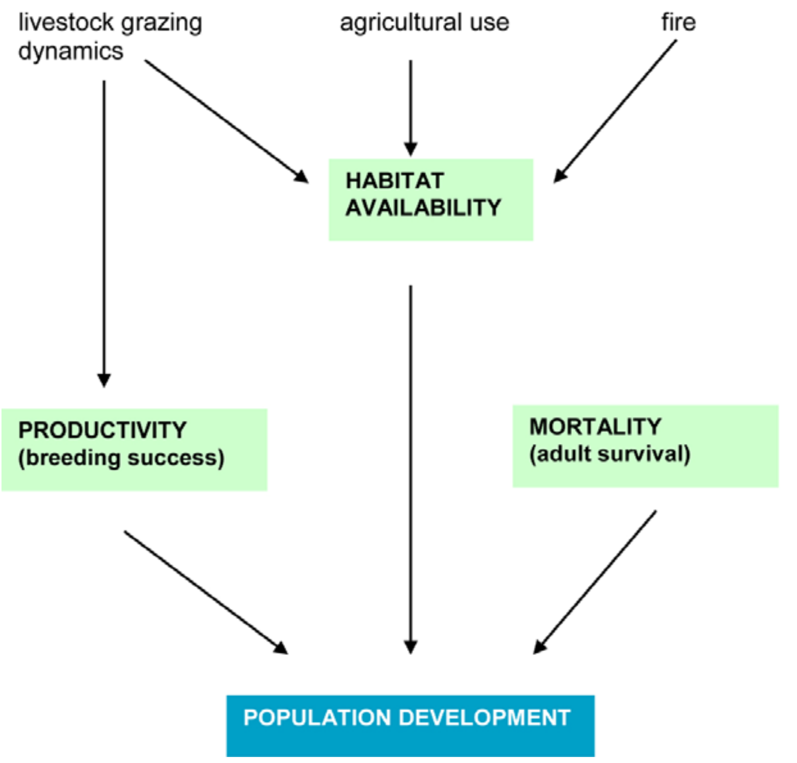


Fig. 20: Factors influencing large scale habitat availability and population development of Sociable Lapwing.

Further reasons for a decline, like the loss of a significant proportion of the population on migration and in the wintering areas due to poaching might contribute as well to the population decline and would increase overall population mortality rates. Thus, there are three superordinated processes which might control Sociable Lapwing population development: The availability of suitable habitat, annual productivity (breeding success) and mortality (Fig. 20).

So far, there has not been a thorough analysis of how habitat availability affected Sociable Lapwing population development. According to my data on the processes influencing habitat selection, I came to the conclusion that this is perhaps the key factor in large scale population trends – probably more influent than changing breeding success or mortality. To preserve Sociable Lapwing (and a suite of steppe species with similar habitat preferences) on a global scale, it is not only necessary to control for breeding success and adult survival, but to monitor socio-geographic processes such as development in agriculture, development of livestock numbers and changes in livestock breeding, and fire fighting.

5.2 Further prospects

This study identified successfully the main factors driving habitat selection of Sociable Lapwing. The results can be used to predict species occurrence in other parts of the breeding range, and to narrow down possible breeding areas in further surveys across Kazakhstan and Russia (planned for 2007 and 2008). The results of this study will be used to identify areas suitable for Sociable Lapwing breeding, and thus to maximize effectiveness in colony search within the huge areas of uniform steppe. For 2007, a first application is planned on 20,000 km² of a new study area in Pavlodar region, NE Kazakhstan. Results of this search can be used to spatially validate my models and test for spatial transferability.

It would be desirable to collect data similar to mine in other regions of the breeding habitat range, as there are hints that breeding habitat use differs across the huge distribution area. By comparing data collected from different regions, models could be refined. However, it is still uncertain if this will be possible in frame of the current research and conservation project.

A very useful approach for further research would include testing, if habitat selected by Sociable Lapwing represents a “functional type” of habitat, and if other species can be added to a suite of steppe species more or less restricted to these areas. There is evidence that especially the vegetation height and distance to human settlements preferred to by Sociable Lapwing are good predictors for the occurrence of breeding Black Lark (*Melanocorypha yeltoniensis*), White-winged Lark (*Melanocorypha leucoptera*) as well. Black- winged Pratincole (*Glareola nordmanni*) and to some extent Caspian Plover (*Charadrius asiaticus*) share the preference for areas close to human settlements situated in the vicinity of water. Even the occurrence of some small mammals and grasshopper species, such as Little Souslik *Spermophilus pygmaeus* and Italian Grasshopper *Calliptamus italicus* seems to be linked with the availability of habitat similar to that preferred by Sociable Lapwing. For all these species there is only very few data available on general and spatial population ecology as well as habitat use and selection. I consider further research on steppe avifauna within the vast grassland range of Northern Kazakhstan most remunerative.

6 ACKNOWLEDGEMENTS

I thank **PD Dr. Robert Biedermann** (Oldenburg, Germany) for agreeing in the subject of this diploma thesis, motivated supervision and help in data analysis with alacrity. **Dr. Robert Sheldon** (London) was a perfect supervisor both in field and during preparation and data analysis. Thanks a lot for having me broadly incorporated into the Sociable Lapwing project for already three years.

Warmest thanks to **Dr. Barbara Strauss** (Oldenburg) for the invaluable support in statistical analysis, and especially for providing self-programmed scripts allowing automated computations and fast data processing in S-Plus.

I want to thank **Dr. Paul Donald** (Sandy, GB), **Steffen Oppel** (Barrow, Alaska), **Holger Schielzeth** (Starnberg, Germany) and **Kilian Wasmer** (Oldenburg) for support and valuable hints concerning sampling design and statistical data analysis.

Maxim A. Koshkin (Astana, Kazakhstan) was most responsible for my fieldwork stays in Kazakhstan being so pleasant. Thanks a lot for all the support.

I am most grateful to **Timur K. Iskakov**, **Kseniya V. Grishina** and **Natal'ya A. Kucheryavaya** for help in data collection and much fun in the field. **Marina Koshkina**, **Vasilii A. Zhulii, Andrew Cotton**, **Andrew Knight**, **Paul Eele**, **John Wills**, **Graeme Ruthven** and **Dave Fairlamb** provided additional assistance during survey work and habitat parameter recording. **Franziska Schwahn** was of invaluable help with plant identification in Kazakhstan. **Tonya A. Grishina** was responsible for excellent accommodation, **Valery N. Grishin** gave important assistance in car repair and maintenance. Thanks to **Tat'yana Siderova**, **Dr. Anna Lushchekina**, **Michael Brombacher, Til Dieterich**, **Aleksei Koshkin**, **Ol'ga Koshkina** and **Prof. Dr. Dietrich Hagen** and for various support and hints.

Great thanks to **Holger Schielzeth, Johanna Maria Karthäuser** and **Kilian Wasmer** for improving earlier drafts of the text by critical reading and helpful discussion.

Finally, warmest thanks to my parents, who not hesitated to support me extensively during the course of my studies both mentally and financially – budgeting far more as usual for literature on ecology and travelling in Eastern Europe and Central Asia. They have a great share in the success of much of my work especially in Ex-USSR countries between 2002 and 2006.

7 ABSTRACTS

7.1 Summary

Due to an assumed sharp decline during the last century and the remaining world population being very small, the Sociable Lapwing (*Vanellus gregarius*) was classified as “Critically Endangered” on the IUCN World Red List in 2005. As part of a long-term research and conservation project, I analyzed habitat and nest site selection of the species in Central Kazakhstan in 2006.

I used a presence/absence modelling approach combined with information theoretic model selection algorithms in order to find the most influent parameters on two hierarchical levels: habitat selection on landscape scale and nest site selection on colony scale.

The results of multivariate logistic regression showed, that on landscape scale, the occurrence probability of Sociable Lapwing is mainly influenced by vegetation height, the density of large grazing animals, and the availability of freshwater. Furthermore, suitable habitat is selected more often close to rivers, which suggests migration following the course of steppe rivers. Topographical and soil parameters were found to be of no influence.

Nest site selection is strongly influenced by the presence of further Sociable Lapwing pairs, as to be expected for a colonial species. Apart from that, I identified a maximised occurrence probability at certain stages of animal dung density and bare ground cover, with 68% of all clutches placed into dung piles (total n=168). Selection preferences for different plant types were not pronounced.

Model fit and discriminative power (evaluated using e.g. Nagelkerke’s R^2, AUC and Cohen’s Kappa) were good to excellent. Internal model validation (bootstrapping procedure) did not impair model performance significantly. Temporal transferability of the results was tested with 2004 to 2007 nest site locations and found to reach 95% correct classification rate, i.e. 95% of all nest sites were found in areas labelled as “suitable” on a habitat suitability map.

Based on these findings, I discuss population development as a function of habitat availability, which fluctuated during the last 100 years due to changing land use practices, population development of wild and domestic ungulates and steppe fire dynamics.

Keywords: Sociable Lapwing, *Vanellus gregarius*, critically endangered, habitat selection, nest site selection, habitat modelling, Central Kazakhstan, steppe, ungulate grazing, *Saiga tatarica*, fire, land use.

7.2 Zusammenfassung

Nach starken Rückgängen im 20. Jahrhundert wurde die Weltpopulation des Steppenkiebitzes Vanellus gregarius im Jahr 2004 auf nur noch 200-600 Brutpaare geschätzt und ein kurzfristiges Aussterben der Art für möglich gehalten. Dies führte zu einer Einstufung der Art als „Critically Endangered“ (kritisch gefährdet) in der weltweiten Roten Liste der IUCN. Als Teilprojekt im Rahmen eines internationalen Forschungs- und Schutzprojektes untersuchte ich die Habitat- und Nistplatzwahl der Art im Sommer 2006 in Zentralkasachstan.

Unter Anwendung eines Präsenz-Absenz-Ansatzes der Habitatmodellierung in Kombination mit informationstheoretischen Methoden zur Modellvereinfachung wurde versucht, die einflussreichsten Parameter im hierarchischen Ablauf der Habitatwahl zu ermitteln.

Die Ergebnisse der Datenanalyse mittels logistischer Regression zeigten, das auf der Landschaftsebene die Antreffwahrscheinlichkeit von Brutkolonien überwiegend von Vegetationshöhe, der Dichte weidender Haustiere und der Verfügbarkeit von Süßwasser abhängig ist. Dies dürften daher auch wesentliche Steuergrößen in der großräumigen Habitatwahl sein. Außerdem wird geeignetes Habitat vorwiegend in der Nähe von Flüssen besetzt. Dies ist möglicherweise mit der Zugstrategie der Art zu erklären. Topographische Parameter wie Höhe über dem Meer, Hangneigung oder –exposition wie auch Bodenparameter hatten keinen Einfluss auf Verteilungsmuster der Kolonien und sind deshalb kaum als erklärende Faktoren in der großräumigen Habitatwahl von Bedeutung.

Wie für eine (halb-)koloniale Watvogelart zu erwarten, beeinflusste die Anwesenheit weiterer Steppenkiebitz-Brutpaare die Nistplatzwahl stark. Weiterhin war die Antreffwahrscheinlichkeit von Nestern innerhalb einer Kolonie an Stellen bestimmter Vegetationsdeckung (etwa 50% offener Boden) und Dungdichte (etwa 20% Deckung) maximiert. 68% aller Nester (n=168) waren in oder dicht neben Dung weidender Haustiere (Pferde, Kühe, Schafe, Ziegen) platziert. Präferenzen für bestimmte Pflanzengesellschaften oder Strukturtypen waren nicht ausgeprägt.

Modellgüte und Modelldiskriminierung wurden unter Zuhilfenahme verschiedener Gütemaße (z.B. Nagelkerke’s R^2, AUC and Cohen’s Kappa) getestet und waren durchwegs gut bis hervorragend. Interne Modellvalidierung mit einem resampling-Verfahren („bootstrapping“) beeinflusste die Modellgüte nicht signifikant. Die zeitliche Übertragbarkeit der erstellten Modelle wurde unter Anwendung auf Brutdaten aus den Jahren 2004, 2005 und 2007 getestet. Etwa 95% der kartierten Neststandorte der Testjahre lagen in Gebieten, die auf Habitateignungskarten als „geeignet“ (mit einer Antreffwahrscheinlichkeit von p>0.5) ausgewiesen worden waren. Dies weist auf eine hervorragende zeitliche Übertragbarkeit der Ergebnisse hin.

Basierend auf den genannten Ergebnissen diskutiere ich die Populationsentwicklung des Steppenkiebitzes als eine Funktion großräumiger Habitatverfügbarkeit. Letztere fluktuierte während der vergangenen 100 Jahre in Abhängigkeit von Landnutzung, Beweidungsmustern wilder und domestizierter Huftiere und Feuerdynamik stark und erklärt beobachtete Populationsschwankungen gut.

Schlüsselwörter: *Steppenkiebitz,* Vanellus gregarius*, critically endangered, Habitatwahl, Nistplatzwahl, Habitatmodellierung, logistische Regression, Kasachstan, Steppe, Beweidung,* Saiga tatarica*, Feuer, Landnutzung.*

7.3 Резюме

В связи с предполагаемым сокращением численности за последнее столетие и очень маленьким размером сохранившейся популяции, в 2005 году кречетка (*Vanellus gregarius*) была внесена в Красный список Международного Союза Охраны Природы (IUCN) как «критически угрожаемый вид». В рамках долгосрочного исследовательского и природоохранного проекта, в 2006 году я проводил анализ выбора местообитания этого вида в Центральном Казахстане.

Я использовал модельный подход присутствия/отсутствия, одновременно с информационной теоретической моделью алгоритмов выбора, для того, чтобы найти наиболее важные параметры на двух иерархических уровнях: выбор местообитания на ландшафтном уровне и выбор места для постройки гнезда на колониальном уровне.

Результаты многомерной логистической регрессии показали, что на ландшафтном уровне возможность присутствия кречетки в основном зависит от высоты растительности, плотности распределения пасущихся животных и доступность пресной воды. К тому же, подходящие места обитания выбираются кречеткой чаще всего вблизи рек, что предполагает ее миграцию вдоль степных рек. Выяснилось, что топографические и почвенные характеристики не оказывают влияния на выбор местообитания.

Выбор места для гнездования очень сильно зависит от присутствия на данной территории других пар кречетки, чего и следует ожидать от колониального вида. Помимо этого, я определил максимальную возможность присутствия, на определенных стадиях, помета крупных животных и участков оголенной почвы, при этом до 68% всех гнезд построены на кучках кизяка. Выбор, зависящий от предпочтения какого-либо типа растений, отмечен не был.

Подбор модели и дискриминационную силу (вычисленные с помощью Nagelkerke's R^2, AUC и Cohen's Kappa) можно классифицировать как «хорошие» и «превосходные». Внутреннее обоснование модели (процедура расчета кривой, используя предыдущее значение для расчета последующего) не повлияло значительно на качество модели. Временная переносимость результатов была проверена на участках размещения гнезд в период с 2004 по 2007 год, и была достигнут 95% уровень точности классификации, т.е. 95 % всех гнезд были найдены на территориях, отмеченных как «подходящие» на карте местообитаний.

Основываясь на полученные данные, я обсуждаю развитие популяции как функцию доступности местообитаний, которая изменялась в течении последних 100 лет в результате перемен в практике землепользования, изменений численности домашних и диких копытных и динамики степных пожаров.

Ключевые слова: кречетка, *Vanellus gregarius*, критически угрожаемый вид, выбор местообитания, статистические модели, логистическая регрессия, Казахстан, степь, выпас скота, землепользование, степные пожары, сайгак, *Saiga tatarica*.

8 REFERENCES

Square brackets [] refer to references in foreign languages (Russian, German).

Akaike, H. 1974. A new look at the statistical model identification. Automatic Control, IEEE Transactions **19**:716-723.

Anderson, D., Burnham, K., Thompson, W. 2000. Null hypothesis testing: Problems, prevalence, and an alternative. Journal of Wildlife Management **64**:912-923.

Atkinson, P. W., Buckingham, D. M., Antony, J. 2004. What factors determine where invertebrate-feeding birds forage in dry agricultural grasslands? Ibis **146**:99-107.

Austin, M. P. 2002. Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. Ecological Modelling **157**:101-118.

Backhaus, K., Erichson, B., Plinke, W., Weiber, R. 2003. [Multivariate statistical analysis - an introduction]. Springer, Berlin.

Barry, S., Elith, J. 2006. Error and uncertainty in habitat models. Journal of Applied Ecology **43**:413–423.

Bekenov, A. B., Grachev, Yu. A., Milner-Gulland, E. J. 1998. The ecology and management of the Saiga antelope in Kazakhstan. Mammal Review **28**:1-52.

Belik, V. P. 1998. [Present situation of the populations of rare and protected waders in southern Russia]. Nesting waders of Eastern Europe, pp. 84-90.

Belik, V. P. 2005. The Sociable Lapwing in Eurasia: what does the future hold? British Birds **98**:76-85.

Belyalov, O. V. 2004. [Sociable Lapwing breeds in Almaty oblast']. Kazakhstanskii Ornithologicheskii Byulleten **3**:152-153.

Berezovikov, N. N., Samusev, I. F., Starikov, S. V. 1998. [The Sociable Lapwing in the Zaissan valley basin and at the Upper Irtysh reaches]. Inform. Materials of the Working Group on Waders **11**:44-47.

Berg, A., T. Lindberg, and K. Kallebrink. 1992. Hatching Success of Lapwings on Farmland:

Differences between Habitats and Colonies of Different Sizes. Journal of Animal Ecology **61**:469-476.

BirdLife International 2006. Sociable Lapwing *Vanellus gregarius*. *In*: IUCN 2006. IUCN Red List of Threatened Species. www.iucnredlist.org.

BirdLife International 2007. Sociable Lapwing: Ornithologists “hit jackpot”. press release at http://www.birdlife.org/news/news/2007/03/sociable_lapwing_discovery.html.

Bonn, A., Schröder, B. 2001. Habitat models and their transfer for single and multi species groups: a case study of carabids in an alluvial forest. Ecography **24**:483-496.

Bragin, E. A. 1999. [Situation of rare bird species in Turgai region and Naurzum reserve, Kazakhstan]. Territorial aspects of bird protection in Central Asia and Kazakhstan, pp. 85-92.

Brotons, L., Thuiller, W., Araujo, M. B., Hirzel, A. H. 2004. Presence-absence versus presence-only modelling methods for predicting bird habitat suitability. Ecography **27**:437-448.

Buckland, S. T., Anderson, D. R., Burnham, K. P., Laake, J. L., Borchers, D. L., Burnham, K., Thomas, L. 2001. Introduction to Distance Sampling: estimating abundance of biological populations. Oxford University Press, USA.

Buckland, S. T., Burnham, K., Augustin, N. H. 1997. Model selection: An integral part of inference. Biometrics **53**:603-618.

Burnham, K. P., Anderson, D. R. 2002. Model selection and multimodel inference: a practical information-theoretic approach. Springer, New York.

Childebaev, M. (ed.) 2003. Environmental monitoring of non-target organisms on the fallow lands of north Kazakhstan during insecticide treatments. IPP Consultants,Technology and institutional development for sustainable locust management. Final Report, Asian development bank program nr. TA No 3647-KAZ.

Cody, M. 1985. Habitat Selection in Birds. Academic Press.

Collar, N. J., Crosby, M. J., Stattersfield, A. J. 1994. Birds to watch 2: The World List of Threatened Birds. BirdLife International, Cambridge.

Cooper, R. J., Dodge, K. M., Martinat, P. J., Donahoe, S. B., Whitmore, R. C. 1990. Effect of diflubenzuron application on eastern deciduous forest birds. Journal of Wildlife

Management **54**:486-493.

Cramp, S., Simmons, K. (eds.) 1983. Handbook of the birds of Europe, the Middle East and North Africa. Vol. 3. Waders to gulls. Oxford University Press, Oxford.

Danchin, E., Boulinier, T., Massot, M. 1998. Conspecific reproductive success and breeding habitat selection: implication for the study of coloniality. Ecology **79**:2415-2428.

Davygora, A. V. 1989. [The Sociable Lapwing *Vanellus gregarius* in the Southern Ural Steppes]. Rare and conservation-dependent animals, pp. 88-90.

De Beurs, K., Henebry, G. 2004. Land surface phenology, climatic variation, and institutional change: analyzing agricultural land cover change in Kazakhstan. Remote Sensing of Environment **89**:497-509.

Dement'ev, G. P., Gladkov, N. A., Spangenberg, E. P. 1951. [Sociable Lapwing *Chettusia gregaria*]. Pp. 101-106 *in*: Dement'ev, G. P., Gladkov, N. A., Spangenberg, E. P., (eds.). Ptitsy sovetskogo soyuza [Birds of the USSR],. Moscow.

Dieterich, T. 2000. [Landscape-ecological research on fallow fields on the territory of the future Biosphere Reserve "Lake Tengiz" in Central Kazakhstan and their possible re-conversion into steppe (sic!)]. Thesis, University of Greifswald, Germany.

Dolgushin, I. A. 1939. [Ornithological observations in Pavlodar oblast’, summer 1939]. Mscr. edited by N. N. Berezovikov and published posthumous (2004) in Transactions of Institute of Zoology **48**:1-39.

Dolgushin, I. A. 1962. [Sociable Lapwing - *Vanellus gregarius*]. Pages 94-101 *in* I. A. Dolgushin, editor. Ptitsy Kazakhstana. Academy of Sciences Kazakh SSR, Alma-Ata.

Doligez, B., Danchin, E., Clobert, J. 2002. Public information and breeding habitat selection in a wild bird population. Science **297**:1168-1170.

Eichhorn, G., Heinicke, T. 2000. Notable observations of the Sociable Lapwing *Vanellus gregarius* from the Tengiz-Korgalzhyn area, Central Kazakhstan. Wader Study Group Bull. **93**:73-76.

Eichhorn, G., Khrokov, V.V. 2002. Decline in breeding Sociable Plover *Chettusia gregaria* in the steppes of Naurzum and Korgalzhyn, Kazakhstan. Sandgrouse **24**:22-27.

Engler, R., Guisan, A., Rechsteiner, R. 2004. An improved approach for predicting the distribution of rare and endangered species from occurrence and pseudo-absence data. Journal of

Applied Ecology **41**:263-274.

Feh, C., Shah, N., Rowen, M., Reading, R., Goyal, S. P. 2002. Status and Action Plan for the Asiatic Wild Ass (*Equus hemionus*). *in* P. D. Moehlman, editor. Equids: Zebras, Asses, and Horses: Status Survey and Conservation Action Plan. IUCN (The World Conservation Union), Gland/Switzerland and Cambridge.

Ferrier, S., Drielsma, M., Manion, G., Watson, G. 2002. Extended statistical approaches to modelling spatial pattern in biodiversity in northeast New South Wales. II. Community-level modelling. Biodiversity and Conservation **11**:2309-2338.

Fielding, A. H., Haworth, P. 1995. Testing the Generality of Bird-Habitat Models. Conservation Biology **9**:1466-1481.

Fielding, A. H., Bell, J. F. 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. Environmental Conservation **24**:38-49.

Finsch, O. 1879. [Expedition to Western Siberia in 1876]. Wallroth, Berlin.

Fisher, R. 1936. The use of multiple measurements in taxonomic problems. Annals of Eugenics **7**:179-188.

Gordienko, N. S. 1991. [Ecology and numbers of Sociable Lapwing in the Kustanai steppe]. Ornitologiya **25**:54-61.

GosKomStat (Russian State Statistics Agency) 2001. Livestock inventories by region, January 1, 1991-2000. *in*. www.ers.usda.gov/Briefing/Russia/Data/LivestockInventoriesByRegion.xls, 23.11.2004.

Grachev, V. A. 1973. [Waders of the Alakol' depression, Eastern Kazakhstan]. Fauna and ecology of waders **2**:28-30.

Graham, M. 2003. Confronting multicollinearity in ecological multiple regression. Ecology **84**:2809-2815.

Greaves, R. K., Sanderson, R. A., Rushton, S. P. 2006. Predicting species occurrence using information-theoretic approaches and significance testing: An example of dormouse distribution in Cumbria, UK. Biological Conservation **130**:239-250.

Guisan, A., Zimmermann, N. 2000. Predictive habitat distribution models in ecology. Ecological Modelling **135**:147-186.

Hanley, J. 1982. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology **143**:29-36.

Harell, F. E. J. 2001. Regression modelling strategies - with applications to linear models, logistic regression, and survival analysis. Springer, New York.

Heinze, G., Schemper, M. 2002. A solution to the problem of separation in logistic regression. Statistics in Medicine **21**:2409-2419.

Hirzel, A., Guisan, A. 2002. Which is the optimal sampling strategy for habitat suitability modelling. Ecological Modelling **157**:331-341.

Hosmer, D., Lemeshow, S. 2000. Applied logistic regression. Wiley, New York.

Ili'chev, V. D., Fomin, B. E. 1979. [The avifauna of Bashkiriya and its changes during the 20th century]. Ornitologiya **14**:83-96.

IPP Consultants 2003. Technology and institutional development for sustainable locust management. Final Report, Asian development bank program nr. TA No 3647-KAZ.

Jeganathan, P., Green, R., Norris, K., Vogiatzakis, I., Bartsch, A., Wotton, S., Bowden, C., Griffiths, G., Pain, D., Rahmani, A. 2004. Modelling habitat selection and distribution of the critically endangered Jerdon's courser *Rhinoptilus bitorquatus* in scrub jungle: an application of a new tracking method. Journal of Applied Ecology **41**:224-237.

Jenness, J. 2004. Nearest features (nearfeat.avx) extension for ArcView 3.x. *in*. Jenness Enterprises, available at: http://www.jennessent.com/arcview/nearest_features.htm.

Jenness, J. 2005. Random point generator (randpts.avx) extension for ArcView 3.x. *in*. Jenness Enterprises, available at: http://www.jennessent.com/arcview/random_points.htm.

Jenness, J. 2005a. Distance Matrix (dist_mat_jen.avx) extension for ArcView 3.x. *in*. Jenness Enterprises, available at: http://www.jennessent.com/arcview/dist_matrix.htm.

Jenness, J. 2005b. Distance/Azimuth Tools (dist_az_tools.avx) extension for ArcView 3.x. *in*. Jenness Enterprises, available at: http://www.jennessent.com/arcview/distance_azimuth.htm.

Johansson, O., Blomqvist, D. 1996. Habitat Selection and Diet of Lapwing *Vanellus vanellus* Chicks on Coastal Farmland in SW Sweden. Journal of Applied Ecology **33**:1030-1040.

Jones, J. 2001. Habitat selection studies in avian ecology: A critical review. Auk **118**:557-562.

Joy, M. K., Death, R. G. 2004. Predicitve modelling and spatial mapping of freshwater fish and decapod assemblages using GIS and neural networks. Freshwater Biology **49**:1046-1052.

Kannan, V. 2004. Birds to watch: Sociable Lapwing *Vanellus gregarius* in India. Birding Asia **1**:78-79.

Karamysheva, Z. V., Rachkovskaya, E. I. 1973. Vegetation of the Central Kazakhstan steppe zone. Academy of Sciences of the USSR, Leningrad.

Karyakin, I. V., Koslov A. A. 1999. [Sociable Lapwing (*Vanellus gregarius*)]. *in*: Karyakin, I.V., Koslov, A. A., (eds.) [Preliminary distribution atlas of the birds of Chelyabinsk province]. Novosibirsk.

Khaidarov, K., Arkhipov, V. 2001. Fire Situation in Kazakhstan. p. 495 *in*: FAO Forestry Department: Global forest fire assessment 1990-2000. Working paper No. 55, United Nations, Rome.

Khrokov, V. V. 1977. [The Sociable Lapwing in the Tengiz-Kurgal'dzhino depression (Central Kazakhstan)]. pp. 231-234 *in* Sludskii, A. A. (ed.): Rare and vanishing animals and birds of Kazakhstan. Alma-Ata.

Khrokov, V. V. 1996. [Breeding biology of Sociable Lapwing *Chettusia gregaria* in Naurzum state nature reserve]. Russkii ornitologicheskii zhurnal **4** (suppl.):3-8.

Khrokov, V. V., Karpov, F.F. 1999. [Population numbers and behaviour of Sociable Lapwing and Little Bustard in Naurzum state nature reserve (Northern Kazakhstan)]. Selevinia **7**:223-225.

Khrokov, V. V. 2000. [Sociable Lapwing (*Vanellus gregarius*) in danger]. Stepnoi bulleten **6**:46-48.

Khrokov, V. V. 2004. [The Sociable Lapwing *Chettusia gregaria* - past and present]. Russkii ornitologicheskii zhurnal **13**:943-961.

Kuchin, F. P., Chekcheev I. P. 1987. [Short notes on Sociable Lapwing at Kulunda lake depression]. Problems of conservation of rare animals, pp. 138-139.

Laing, S. E., Buckland, S. T., Burn, R. W., Lambie, D., Amphlett, A. 2003. Dung and nest surveys: estimating decay rates. Journal of Applied Ecology **40**:1102-1111.

Langgemach, T., Bellebaum, J. 2005. [Predation and conservation of ground-breeding birds in Germany]. Vogelwelt **126**:259-294.

Lenk, M. 2001. [Socio-economic transformation processes in rural Central Kazakhstan after the country's independence in 1991]. Thesis, University of Greifswald, Germany.

López-López, P., Garcia-Ripollés, C., Aguilar, J., Garcia-López, F., Verdejo, J. 2006. Modelling breeding habitat preferences of Bonelli's eagle (*Hieraaetus fasciatus*) in relation to topography, disturbance, climate and land use at different spatial scales. Journal of Ornithology **147**:97-106.

MacKenzie, D., Nichols, J., Lachman, G., Droege, S., Royle, J., Langtimm, C. 2002. Estimating site occupancy rates when detection rates are less than one. Ecology **83**:2248-2255.

Malovichko, L. V., Fedosov, V. N., Kurochkin, E. N., Eltyshev, S. D., Slin'ko, A. V. 2006. [New data on Sociable Lapwing *Chettusia gregaria* in Stavropol Region, Northern Caucasus, Russia]. Information Materials of the Russian Wader Study Group **19**:45-47.

Manel, S., Dias, J., Ormerod, J. 1999. Comparing discriminant analysis, neural networks and logistic regression for predicting species distributions: a case study with a Himalayan river bird. Ecological Modelling **120**:337-347.

Manel, S., Williams, H.C., Ormerod, S. J. 2001. Evaluating presence-absence models in ecology: the need to account for prevalence. Journal of Applied Ecology **38**:921-931.

Manly, B. F. J. 2001. Randomization, bootstrap and Monte Carlo methods in biology. Chapman and Hall, London.

Martin, T. 1987. Food as a Limit on Breeding Birds: A Life-History Perspective. Annual Review of Ecology and Systematics **18**:453-487.

McNeil, B., Hanley, J. 1984. Statistical approaches to the analysis of receiver operating characteristic (ROC) curves. Med Decis Making **4**:137-150.

Meng, E., Longmire, J., Moldashev, A. 2000. Kazakhstan's wheat system: priorities, constraints, and future prospects. Food Policy **25**:701-717.

Michalek, K. G., Auer, J. A., Großberger, H., Schmalzer, A., Winkler, H. 2001. [Influence of habitat, weather and forest management on breeding density of Great Spotted Woodpeckers and Middle Spotted Woodpeckers (*Picoides major and P. medius*) in the Vienna Woods

("Wienerwald")]. Abhandl. Ber. Mus. Heineanum **5**:31-58.

Milner-Gulland, E. J., Kholodova, M. V., Bekenov, A., Bukreeva, O. M., Grachev, Iu., Amgalan, L., Lushchekina, A.A. 2001. Dramatic declines in saiga antelope populations. Oryx **35**:340-345.

Monserud, R. A., Leemans, R. 1992. Comparing Global Vegetation Maps with the Kappa-Statistic. Ecological Modelling **62**:275-293.

Mooney, C., Duval, R. 1993. Bootstrapping: A Nonparametric Approach to Statistical Inference. Sage, Newbury Park.

Morozov, V. V. 2005. Sociable Plover surveys in steppes of Southern Russia 2005. Final report on a pilot project. BirdLife International, Moscow.

Mullie, W. C., Keith, J. O. 1993. The effects of aerially applied Fenitrothion and chlorpyriphos on birds in the savannah of northern Senegal. Journal of Applied Ecology **30**:536-550.

Nagelkerke, N. 1991. A note on a general definition of the coefficient of determination. Biometrika **78**:691-692.

Olson, D. M., Dinerstein, E., Wikramanayake, E. D., Burgess, N. D., Powell, G. V. N., Underwood, E. C., D'Amico, J. A., Itoua, I., Strand, H. E., Morrison, J. C., Loucks, C. J., Allnutt, T. F., Ricketts, T. H., Kura, Y., Lamoreux, J. F., Wettengel, W. W., Hedao, P., Kassem, K. R. 2001. Terrestrial Ecoregions of the World: A New Map of Life on Earth. BioScience **51**:933-938.

Oppel, S., Schaefer, H. M., Schmidt, V., Schröder, B. 2004. Habitat selection by the pale-headed brush-finch (*Atlapetes pallidiceps*) in southern Ecuador: implications for conservation. Biological Conservation **118**:33-40.

Osborne, P. E., Alonso, J. C., Bryant, R. G. 2001. Modelling landscape-scale habitat using GIS and remote sensing: a case study with great bustards. Journal of Applied Ecology **38**:458-471.

Özesmi, S., Özesmi, U. 1999. An artificial neural network approach to spatial habitat modelling with interspecific interaction. Ecological Modelling **116**:15-31.

Pearce, J., Ferrier, S. 2001. The practical value of modelling relative abundance of species for regional conservation planning: a case study. Biological Conservation **98**:33-43.

Posada, D., Buckley, T. R. 2004. Model selection and model averaging in phylogenetics:

Advantages of Akaike information criterion and Bayesian approaches over likelihood ratio tests. Systematic Biology **53**:793-808.

Prasolova, L. I., Gerasimov, I. P. (eds.). 1955. [State soil map of the USSR 1:1 Mio. Sheet M 42 - Akmolinsk]. Academy of Sciences of the USSR., Moscow.

Radde, G. 1884. [Ornis Caucasica: A caucasian avifauna]. Fischer, Kassel.

Redkov, V. V. 1964. [Soils of the Tselinograd Oblast']. Nauka, Alma-Ata.

Redmann, R., Romo, J., Pylypec, B., Driver, E. 1993. Impacts of Burning on Primary Productivity of *Festuca* and *Stipa-Agropyron* Grasslands in Central Saskatchewan. American Midland Naturalist **130**:262-273.

Reineking, B., Schröder, B. 2004. [Variable selection in statistical modelling]. *In*: Dormann, C. F., Blaschke, T., Lausch, A., Schröder, B., Söndgerath, D., (eds.)[Habitat modelling - methods, application, benefits]. UFZ, Leipzig/Halle.

Robinson, S., Milner-Gulland, E. J. 2003. Political change and factors limiting numbers of wild and domestic ungulates in Kazakhstan. Human Ecology **31**:87-110.

Ryabov, V. F. 1949. [On the ecology of some steppe birds of Northern Kazakhstan based on observations in the Naurzum zapovednik]. Trudy Naurzumskogo Zapovednika **3**:12-16.

Ryabov, V. F. 1974. [Changes in bird communities of the steppes of Northern Kazakhstan under human influence]. Ornitologiya **11**:279-297.

Samigulin, G. M. 1987. [Short notes on Sociable Lapwing *Vanellus gregarius* in Orenburg region, Russia]. Problems in the conservation of rare and vanishing animals, pp. 137-138.

Sample, B. E., Cooper, R. J., Whitmore, R. C. 1993. Dietary shifts among songbirds from a diflubenzuron-treated forest. Condor **95**:616-624.

Schielzeth, H. 2005. [The Sociable Lapwing - an endangered endemic of the Eurasian Steppe. Results of a research project in Central Asia]. Falke **52**:44-49.

Schielzeth, H., Eichhorn, G., Heinecke, T., Kamp, J., Koshkin, M.A., Koshkin A.V., Lachmann, L. (2007, in press): Waterbird population estimates for a key staging site in Kazakhstan – a contribution to wetland conservation on the Central Asian flyway. Bird Conservation International 17:xx-xx.

Schröder, B., Reineking, B. 2004. [Modelling species-area-relationship - an overview of mehtods currently used in habitat modelling]. *in*: Dormann, C. F., Blaschke, T., Lausch, A., Schröder, B., Söndgerath, D., editor. [Habitat modelling - methods, application, benefits]. UFZ, Leipzig/Halle.

Schwahn, F. 2006. Plants and vegetation of fallow fields in Korgalzhyn region, Central Kazakhstan. An illustrated guide, CD-ROM.

Sharleman', M. V. 1938. [Birds of the Ukrainian SSR]. Acad Sci Ukr SSR, Kiev.

Sheldon, R., Koshkin, M. A., Kamp, J., Khrokov, V. V., Donald, P. 2005. Sociable Lapwing *Vanellus gregarius* project - preliminary fieldwork report 2005. RSPB, Sandy, UK.

Sheldon, R., Grishina, K., Kamp, J., Khrokov, V., Knight, A., Koshkin, M. 2006. Revising the breeding population estimate and distribution of the Critically Endangered Sociable Lapwing *Vanellus gregarius*. p. 77 *in*: Sitters, H. (ed.). Proceedings of the 12th International Wader Study Group Annual Conference, Falsterbo, Sweden.

Shevchenko, V. L., Debelo, P. V., Gavrilov, E. I., Naglov, V. A., Fedosenko, A. K. 1993. [About the avifauna of the Volga-Ural interfluve]. *in* E. I. Gavrilov, editor. Fauna and biology of birds of the Kazakhstan [sic]. Nauka, Almaty.

Shevchenko, V. L. 1998. [Current situation of the Sociable Plover north of the Caspian Sea]. Inform. Materials of the Russian Wader Study Group **11**:45-47.

Siderova, T. V. 1988. [Vascular plants of the Kurgald'zhinksii Zapovednik]. Moscow.

Sokolov, V. E., Zhirnov, L. V. 1998. [Saiga antelope: phylogeny, systematics, ecology, conservation, and use]. Russian Academy of Sciences, Moscow.

Solomatin, A. O. 1997. [The Sociable Lapwing at the Irtysh valley near Pavlodar, Kazakhstan]. Bulleten’ Mosk. Obshestva Ispytatelei Prirody otdel Biologii **12**:23-28.

Strauss, B., Biedermann, R. 2006. Urban brownfields as temporary habitats: Driving forces for the diversity of phytophagous insects. Ecography **29**:928-940.

Suleimenov, M., Oram, P. 2000. Trends in feed, livestock production, and rangelands during the transition period in three Central Asian countries. Food Policy **25**:681-700.

Sushkin, P. P. 1908. [The birds of the central Khyrgyz steppe]. Russian Academy of Sciences, Moscow.

Sutherland, W. J., Newton, I., Green, R. 2004. Bird Ecology and Conservation - A Handbook of Techniques. Oxford University Press, Oxford.

Teunissen, W. A., Schekkermann, H., van Passen, A. 2005. [Meadowbirds and Predation. Newsletter of Meadowbirds and Predation project, 2004 season]. SOVON, Beek-Ubbergen, The Netherlands.

Tomkovich, P. S., Lebedeva, E. A. 2004. International Single Species Action Plan for the conservation of the Sociable Lapwing *Vanellus gregarius*. AEWA technical series No.2, AEWA/BirdLife International, Moscow.

Tucker, G. M. 1992. Effects of agricultural practices on field use by invertebrate-feeding birds in winter. J. Appl. Ecol. 29:779-790.

Tyre, A., Tenhumberg, B., Field, S., Niejalke, D., Parris, K., Possingham, H. 2003. Improving precision and reducing bias in biological surveys: estimating false-negative error rates. Ecological Applications **13**:1790-1801.

Volchanetskii, V. K. 1937. [About the ornithofauna of the Volga-Ural steppes]. Transactions of the Khar'kov Scientific Research Institute, pp. 21-81.

Vorob'ev, K. A. 1936. [Materials to the ornithofauna of the Volga delta and adjacent steppe regions]. Transactions of the Astrakhan' Reserve, pp. 7-60.

Watson, M., Wilson, J. M., Koshkin, M., Sherbakov, B., Karpov, F., Gavrilov, A., Schielzeth, H., Brombacher, M., Collar, N. J., Cresswell, W. 2006. Nest survival and productivity of the critically endangered Sociable Lapwing *Vanellus gregarius*. Ibis **148**:489-502.

Wein, N. 1983. [The Soviet Union]. UTB/Schoeningh, Paderborn.

Whitmore, R. C., Sample, B. E., Cooper, R. J. 1993. Bird fat levels as a measure of effect in forests treated with diflubenzuron. Environmental Toxicology and Chemistry **12**:2059-2064.

Whittingham, M. J., Stephens, P. A., Bradbury, R. B., Freckleton, R. P. 2006. Why do we still use stepwise modelling in ecology and behaviour? Journal of Animal Ecology **75**:1182–1189.

Wright, I. A. 2002. Livestock production and rural development in Kazakhstan. Ann. report Macaulay Institute **12**:39-41.

ZelinGiProZem. 1986. State farm "Amangel'dy". Map 1:50,000. Tselinograd.

ZelinGiProZem. 1990. State farm "Druzhba". Map 1:50,000. Tselinograd.

ZelinGiProZem. 1991. State farm "Kurgald'zhinskii". Map 1:50,000. Tselinograd.

EIDESSTATTLICHE ERKLÄRUNG

Ich versichere an Eides Statt, dass ich die von mir eingereichte Diplomarbeit in sämtlichen Teilen selbständig verfasst und ausschließlich die angegebenen Hilfsmittel und Quellen benutzt habe.
Die Arbeit wurde bisher in gleicher oder ähnlicher Form keiner anderen Prüfungsbehörde vorgelegt und nicht veröffentlicht.

Oldenburg, den 14.04.2007

Johannes Kamp
Hindenburgstr. 3
D- 26122 Oldenburg
Germany

johannes.kamp@uni-oldenburg.de
johannes_kamp@mail.ru

APPENDICES

APPENDICES

Appendix A. Supplementary analysis material

Appendix B. Scripts used for automated model calculations in S-Plus

Appendix C. Maps

Appendix A. Supplementary analysis material

App. A-1. Correlation analysis for all variables considered at the landscape scale (Spearman's rank correlation test). For variable shortcuts cf. chapter 2.2.4, Table 3.

		elev	slope	asp	cov.dung	cov.art	cov.fest	cov.stip	cov.gras	cov.ML	cov.herb	cov.soil	cov.veg	vegH.mod	vegH.max	dist.col	dist.lake	dist.riv	dist.wat	dist.rook	dist.sett	dung.fre	dung.old	dung.tot
Spearman's r	**elev**	1.000	0.400	-0.399	-0.205	0.005	0.133	0.209	0.002	0.076	0.047	0.149	-0.120	-0.008	-0.040	0.406	0.093	0.161	0.228	0.357	0.464	-0.128	-0.286	-0.283
significance (2-sided)		.	0.000	0.000	0.015	0.952	0.119	0.013	0.977	0.372	0.583	0.080	0.158	0.923	0.640	0.000	0.275	0.058	0.007	0.000	0.000	0.132	0.001	0.001
Spearman's r	**slope**	0.400	1.000	-0.955	-0.049	-0.053	0.181	0.222	0.063	0.266	0.045	0.090	-0.090	-0.032	-0.034	0.112	0.085	0.102	0.146	0.047	0.241	-0.092	-0.054	-0.081
significance (2-sided)		0.000	.	0.000	0.568	0.533	0.033	0.009	0.464	0.002	0.599	0.293	0.294	0.712	0.692	0.188	0.321	0.230	0.086	0.586	0.004	0.279	0.525	0.341
Spearman's r	**asp**	-0.399	-0.955	1.000	0.018	0.012	-0.194	-0.216	-0.073	-0.248	-0.039	-0.076	0.077	0.037	0.035	-0.105	-0.072	-0.144	-0.155	-0.048	-0.223	0.068	0.033	0.057
significance (2-sided)		0.000	0.000	.	0.835	0.889	0.022	0.011	0.395	0.003	0.648	0.374	0.367	0.669	0.683	0.216	0.401	0.092	0.068	0.578	0.008	0.423	0.700	0.502
Spearman's r	**cov.dung**	-0.205	-0.049	0.018	1.000	0.141	0.157	-0.118	-0.062	-0.030	-0.165	0.051	-0.146	-0.393	-0.287	-0.283	0.021	-0.446	-0.344	-0.404	-0.576	0.607	0.704	0.710
significance (2-sided)		0.015	0.568	0.835	.	0.097	0.065	0.166	0.466	0.722	0.052	0.549	0.087	0.000	0.001	0.001	0.807	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Spearman's r	**cov.art**	0.005	-0.053	0.012	0.141	1.000	-0.201	-0.231	-0.365	0.189	-0.282	-0.030	0.026	-0.107	-0.205	-0.085	0.180	-0.215	-0.013	-0.119	-0.263	0.248	0.171	0.203
significance (2-sided)		0.952	0.533	0.889	0.097	.	0.017	0.006	0.000	0.026	0.001	0.725	0.764	0.209	0.015	0.318	0.034	0.011	0.875	0.163	0.002	0.003	0.044	0.017
Spearman's r	**cov.fest**	0.133	0.181	-0.194	0.157	-0.201	1.000	0.492	0.576	0.276	0.134	-0.206	0.197	0.074	0.265	-0.006	0.012	0.093	0.158	-0.026	0.138	-0.021	0.111	0.077
significance (2-sided)		0.119	0.033	0.022	0.065	0.017	.	0.000	0.000	0.001	0.116	0.015	0.020	0.388	0.002	0.940	0.890	0.276	0.063	0.761	0.106	0.805	0.195	0.366
Spearman's r	**cov.stip**	0.209	0.222	-0.216	-0.118	-0.231	0.492	1.000	0.591	0.117	0.194	-0.295	0.308	0.350	0.516	0.088	0.113	0.141	0.208	0.096	0.257	-0.240	-0.122	-0.149
significance (2-sided)		0.013	0.009	0.011	0.166	0.006	0.000	.	0.000	0.169	0.022	0.000	0.000	0.000	0.000	0.302	0.187	0.097	0.014	0.259	0.002	0.005	0.152	0.080
Spearman's r	**cov.grass**	0.002	0.063	-0.073	-0.062	-0.365	0.576	0.591	1.000	0.006	0.052	-0.552	0.555	0.489	0.691	-0.120	0.037	0.074	0.128	-0.128	0.080	-0.061	0.045	0.031
significance (2-sided)		0.977	0.464	0.395	0.466	0.000	0.000	0.000	.	0.942	0.542	0.000	0.000	0.000	0.000	0.159	0.666	0.389	0.133	0.132	0.351	0.473	0.600	0.721
Spearman's r	**cov.ML**	0.076	0.266	-0.248	-0.030	0.189	0.276	0.117	0.006	1.000	0.076	-0.174	0.174	-0.028	-0.030	-0.022	0.103	0.057	0.105	0.053	0.040	-0.024	0.109	0.118
significance (2-sided)		0.372	0.002	0.003	0.722	0.026	0.001	0.169	0.942	.	0.375	0.041	0.040	0.742	0.724	0.797	0.229	0.503	0.219	0.539	0.642	0.782	0.201	0.166
Spearman's r	**cov.herb**	0.047	0.045	-0.039	-0.165	-0.282	0.134	0.194	0.052	0.076	1.000	-0.278	0.283	0.322	0.229	0.101	0.066	0.152	0.270	0.132	0.344	-0.242	-0.234	-0.251
significance (2-sided)		0.583	0.599	0.648	0.052	0.001	0.116	0.022	0.542	0.375	.	0.001	0.001	0.000	0.007	0.236	0.441	0.075	0.001	0.121	0.000	0.004	0.006	0.003
Spearman's r	**cov.soil**	0.149	0.090	-0.076	0.051	-0.030	-0.206	-0.295	-0.552	-0.174	-0.278	1.000	-0.990	-0.585	-0.623	0.103	-0.094	-0.006	-0.122	0.168	-0.039	0.005	-0.015	-0.031
significance (2-sided)		0.080	0.293	0.374	0.549	0.725	0.015	0.000	0.000	0.041	0.001	.	0.000	0.000	0.000	0.227	0.273	0.940	0.154	0.048	0.646	0.956	0.857	0.714
Spearman's r	**cov.veg**	-0.120	-0.090	0.077	-0.146	0.026	0.197	0.308	0.555	0.174	0.283	-0.990	1.000	0.623	0.647	-0.089	0.102	0.045	0.153	-0.129	0.092	-0.068	-0.050	-0.035
significance (2-sided)		0.158	0.294	0.367	0.087	0.764	0.020	0.000	0.000	0.040	0.001	0.000	.	0.000	0.000	0.298	0.231	0.599	0.073	0.130	0.282	0.428	0.560	0.684
Spearman's r	**vegH.mod**	-0.008	-0.032	0.037	-0.393	-0.107	0.074	0.350	0.489	-0.028	0.322	-0.585	0.623	1.000	0.829	0.031	0.107	0.152	0.223	0.007	0.347	-0.338	-0.360	-0.376
significance (2-sided)		0.923	0.712	0.669	0.000	0.209	0.388	0.000	0.000	0.742	0.000	0.000	0.000	.	0.000	0.720	0.210	0.075	0.008	0.935	0.000	0.000	0.000	0.000
Spearman's r	**vegH.max**	-0.040	-0.034	0.035	-0.287	-0.205	0.265	0.516	0.691	-0.030	0.229	-0.623	0.647	0.829	1.000	-0.030	0.101	0.063	0.203	-0.046	0.260	-0.276	-0.207	-0.226
significance (2-sided)		0.640	0.692	0.683	0.001	0.015	0.002	0.000	0.000	0.724	0.007	0.000	0.000	0.000	.	0.729	0.237	0.460	0.016	0.589	0.002	0.001	0.015	0.008
Spearman's r	**dist.col**	0.406	0.112	-0.105	-0.283	-0.085	-0.006	0.088	-0.120	-0.022	0.101	0.103	-0.089	0.031	-0.030	1.000	-0.153	0.443	0.184	0.541	0.592	-0.309	-0.404	-0.402
significance (2-sided)		0.000	0.188	0.216	0.001	0.318	0.940	0.302	0.159	0.797	0.236	0.227	0.298	0.720	0.729	.	0.072	0.000	0.030	0.000	0.000	0.000	0.000	0.000
Spearman's r	**dist.lake**	0.093	0.085	-0.072	0.021	0.180	0.012	0.113	0.037	0.103	0.066	-0.094	0.102	0.107	0.101	-0.153	1.000	-0.206	0.519	-0.077	0.115	0.010	-0.056	-0.041
significance (2-sided)		0.275	0.321	0.401	0.807	0.034	0.890	0.187	0.666	0.229	0.441	0.273	0.231	0.210	0.237	0.072	.	0.015	0.000	0.371	0.177	0.910	0.514	0.631
Spearman's r	**dist.riv**	0.161	0.102	-0.144	-0.446	-0.215	0.093	0.141	0.074	0.057	0.152	-0.006	0.045	0.152	0.063	0.443	-0.206	1.000	0.493	0.471	0.459	-0.410	-0.439	-0.462

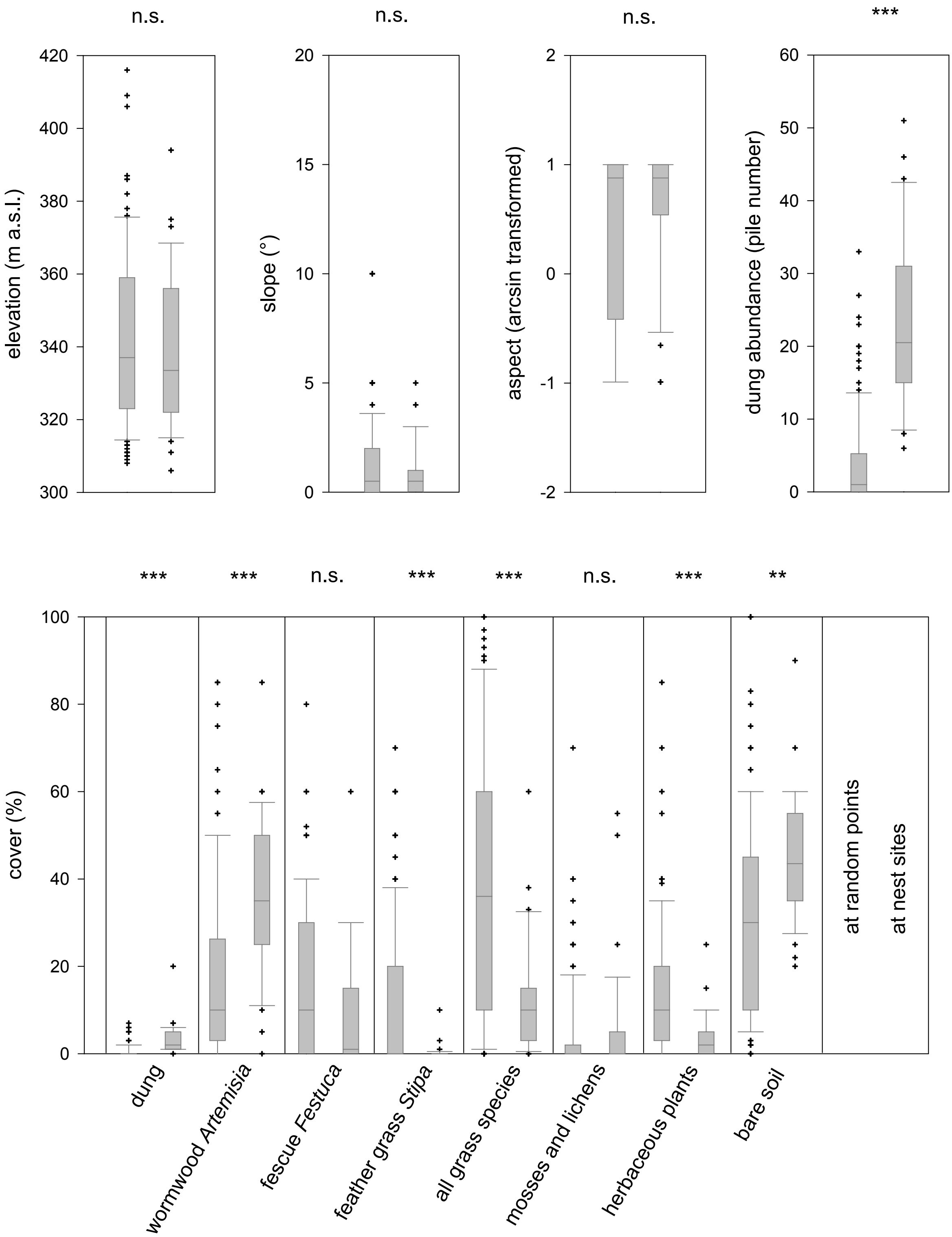
significance (2-sided)		0.058	0.230	0.092	0.000	0.011	0.276	0.097	0.389	0.503	0.075	0.940	0.599	0.075	0.460	0.000	0.015	.	0.000	0.000	0.000	0.000	0.000	0.000
Spearman's r	**dist.wat**	0.228	0.146	-0.155	-0.344	-0.013	0.158	0.208	0.128	0.105	0.270	-0.122	0.153	0.223	0.203	0.184	0.519	0.493	1.000	0.188	0.483	-0.260	-0.361	-0.365
significance (2-sided)		0.007	0.086	0.068	0.000	0.875	0.063	0.014	0.133	0.219	0.001	0.154	0.073	0.008	0.016	0.030	0.000	0.000	.	0.027	0.000	0.002	0.000	0.000
Spearman's r	**dist.rook**	0.357	0.047	-0.048	-0.404	-0.119	-0.026	0.096	-0.128	0.053	0.132	0.168	-0.129	0.007	-0.046	0.541	-0.077	0.471	0.188	1.000	0.397	-0.300	-0.507	-0.493
significance (2-sided)		0.000	0.586	0.578	0.000	0.163	0.761	0.259	0.132	0.539	0.121	0.048	0.130	0.935	0.589	0.000	0.371	0.000	0.027	.	0.000	0.000	0.000	0.000
Spearman's r	**dist.sett**	0.464	0.241	-0.223	-0.576	-0.263	0.138	0.257	0.080	0.040	0.344	-0.039	0.092	0.347	0.260	0.592	0.115	0.459	0.483	0.397	1.000	-0.540	-0.692	-0.715
significance (2-sided)		0.000	0.004	0.008	0.000	0.002	0.106	0.002	0.351	0.642	0.000	0.646	0.282	0.000	0.002	0.000	0.177	0.000	0.000	0.000	.	0.000	0.000	0.000
Spearman's r	**dung.fre**	-0.128	-0.092	0.068	0.607	0.248	-0.021	-0.240	-0.061	-0.024	-0.242	0.005	-0.068	-0.338	-0.276	-0.309	0.010	-0.410	-0.260	-0.300	-0.540	1.000	0.559	0.665
significance (2-sided)		0.132	0.279	0.423	0.000	0.003	0.805	0.005	0.473	0.782	0.004	0.956	0.428	0.000	0.001	0.000	0.910	0.000	0.002	0.000	0.000	.	0.000	0.000
Spearman's r	**dung.old**	-0.286	-0.054	0.033	0.704	0.171	0.111	-0.122	0.045	0.109	-0.234	-0.015	-0.050	-0.360	-0.207	-0.404	-0.056	-0.439	-0.361	-0.507	-0.692	0.559	1.000	0.980
significance (2-sided)		0.001	0.525	0.700	0.000	0.044	0.195	0.152	0.600	0.201	0.006	0.857	0.560	0.000	0.015	0.000	0.514	0.000	0.000	0.000	0.000	0.000	.	0.000
Spearman's r	**dung.tot**	-0.283	-0.081	0.057	0.710	0.203	0.077	-0.149	0.031	0.118	-0.251	-0.031	-0.035	-0.376	-0.226	-0.402	-0.041	-0.462	-0.365	-0.493	-0.715	0.665	0.980	1.000
significance (2-sided)		0.001	0.341	0.502	0.000	0.017	0.366	0.080	0.721	0.166	0.003	0.714	0.684	0.000	0.008	0.000	0.631	0.000	0.000	0.000	0.000	0.000	0.000	.

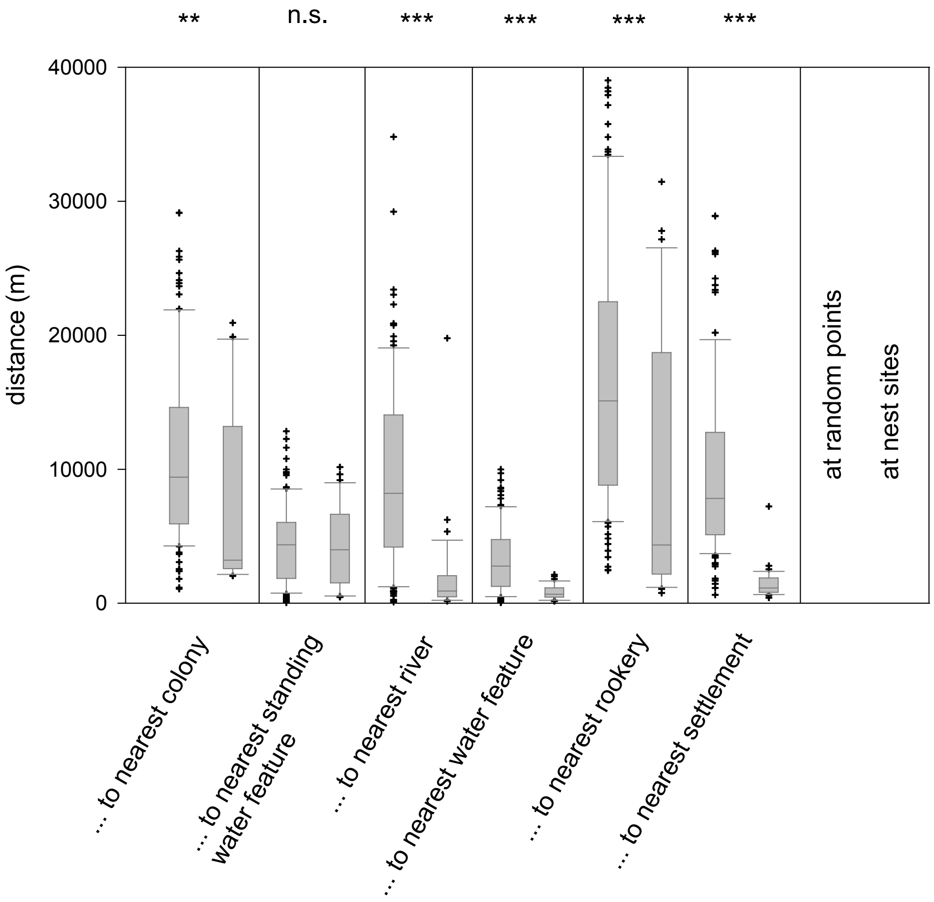
App. A-2. Correlation analysis for all variables considered at the colony scale (Spearman's rank correlation test). For variable shortcuts cf. chapter 2.2.4, Table 3.

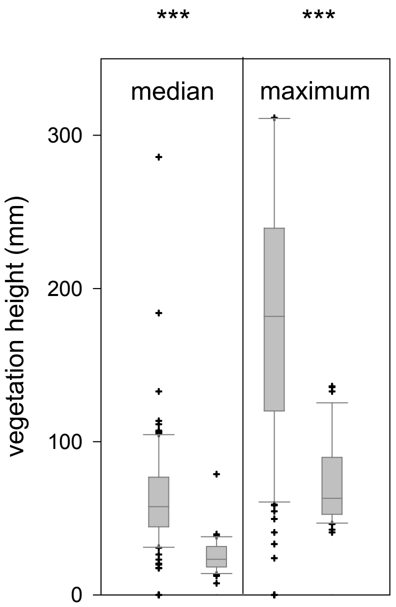
		cov.dung	**cov.Art**	**cov.Fest**	**cov.Stip**	**cov.grass**	**cov.ML**	**cov.herb**	**cov.soil**	**cov.veg**	**vegH.mod**	**vegH.max**	**dist.nest**	**dist.lake**	**dist.water**	**dist.riv**	**dist.rook**	**dist.sett**
Spearman's r	**cov.dung**	1.000	-0.025	0.161	-0.069	-0.133	0.197	-0.125	0.043	-0.155	-0.125	-0.145	0.068	-0.113	-0.069	-0.065	-0.006	-0.035
significance (2-sided)		.	0.639	0.003	0.200	0.014	0.000	0.020	0.432	0.004	0.020	0.007	0.209	0.036	0.200	0.226	0.914	0.519
Spearman's r	**cov.Art**	-0.025	1.000	-0.148	-0.096	-0.342	-0.119	-0.278	-0.024	0.026	-0.152	-0.172	0.061	-0.033	0.104	0.043	-0.073	-0.012
significance (2-sided)		0.639	.	0.006	0.075	0.000	0.027	0.000	0.663	0.629	0.005	0.001	0.261	0.548	0.054	0.432	0.178	0.819
Spearman's r	**cov.Fest**	0.161	-0.148	1.000	0.241	-0.244	0.426	-0.241	-0.302	0.281	-0.077	-0.079	-0.025	0.010	0.011	-0.106	0.058	0.005
significance (2-sided)		0.003	0.006	.	0.000	0.000	0.000	0.000	0.000	0.000	0.153	0.145	0.646	0.853	0.834	0.049	0.280	0.925
Spearman's r	**cov.Stip**	-0.069	-0.096	0.241	1.000	-0.130	0.150	-0.057	-0.222	0.228	-0.009	0.020	0.006	0.196	-0.036	-0.079	-0.057	0.012
significance (2-sided)		0.200	0.075	0.000	.	0.016	0.005	0.296	0.000	0.000	0.866	0.710	0.915	0.000	0.510	0.147	0.289	0.831
Spearman's r	**cov.grass**	-0.133	-0.342	-0.244	-0.130	1.000	-0.147	-0.101	-0.430	0.440	0.116	0.139	0.007	-0.072	-0.131	-0.021	0.078	0.142
significance (2-sided)		0.014	0.000	0.000	0.016	.	0.006	0.063	0.000	0.000	0.032	0.010	0.891	0.185	0.015	0.702	0.152	0.008
Spearman's r	**cov.ML**	0.197	-0.119	0.426	0.150	-0.147	1.000	-0.138	-0.295	0.269	-0.054	-0.055	-0.048	-0.051	-0.066	-0.061	0.066	0.158
significance (2-sided)		0.000	0.027	0.000	0.005	0.006	.	0.011	0.000	0.000	0.316	0.307	0.378	0.348	0.220	0.263	0.225	0.003
Spearman's r	**cov.herb**	-0.125	-0.278	-0.241	-0.057	-0.101	-0.138	1.000	-0.186	0.198	0.482	0.501	-0.050	-0.007	-0.053	0.109	0.009	-0.017
significance (2-sided)		0.020	0.000	0.000	0.296	0.063	0.011	.	0.001	0.000	0.000	0.000	0.353	0.902	0.331	0.044	0.866	0.755
Spearman's r	**cov.soil**	0.043	-0.024	-0.302	-0.222	-0.430	-0.295	-0.186	1.000	-0.994	-0.237	-0.276	0.003	0.069	0.064	0.006	-0.082	-0.165
significance (2-sided)		0.432	0.663	0.000	0.000	0.000	0.000	0.001	.	0.000	0.000	0.000	0.957	0.201	0.235	0.913	0.130	0.002
Spearman's r	**cov.veg**	-0.155	0.026	0.281	0.228	0.440	0.269	0.198	-0.994	1.000	0.249	0.289	-0.011	-0.056	-0.056	0.002	0.082	0.168
significance (2-sided)		0.004	0.629	0.000	0.000	0.000	0.000	0.000	0.000	.	0.000	0.000	0.845	0.304	0.303	0.978	0.131	0.002
Spearman's r	**vegH.mod**	-0.125	-0.152	-0.077	-0.009	0.116	-0.054	0.482	-0.237	0.249	1.000	0.958	0.002	0.004	-0.115	-0.021	-0.048	0.016
significance (2-sided)		0.020	0.005	0.153	0.866	0.032	0.316	0.000	0.000	0.000	.	0.000	0.971	0.941	0.033	0.698	0.380	0.765
Spearman's r	**vegH.max**	-0.145	-0.172	-0.079	0.020	0.139	-0.055	0.501	-0.276	0.289	0.958	1.000	0.005	0.004	-0.126	-0.014	-0.036	0.050
significance (2-sided)		0.007	0.001	0.145	0.710	0.010	0.307	0.000	0.000	0.000	0.000	.	0.934	0.948	0.019	0.799	0.506	0.357
Spearman's r	**dist.nest**	0.068	0.061	-0.025	0.006	0.007	-0.048	-0.050	0.003	-0.011	0.002	0.005	1.000	-0.010	0.017	-0.028	0.029	-0.092
significance (2-sided)		0.209	0.261	0.646	0.915	0.891	0.378	0.353	0.957	0.845	0.971	0.934	.	0.847	0.752	0.602	0.587	0.089
Spearman's r	**dist.lake**	-0.113	-0.033	0.010	0.196	-0.072	-0.051	-0.007	0.069	-0.056	0.004	0.004	-0.010	1.000	-0.187	-0.395	-0.226	-0.197
significance (2-sided)		0.036	0.548	0.853	0.000	0.185	0.348	0.902	0.201	0.304	0.941	0.948	0.847	.	0.001	0.000	0.000	0.000
Spearman's r	**dist.water**	-0.069	0.104	0.011	-0.036	-0.131	-0.066	-0.053	0.064	-0.056	-0.115	-0.126	0.017	-0.187	1.000	0.333	0.434	-0.221
significance (2-sided)		0.200	0.054	0.834	0.510	0.015	0.220	0.331	0.235	0.303	0.033	0.019	0.752	0.001	.	0.000	0.000	0.000
Spearman's r	**dist.riv**	-0.065	0.043	-0.106	-0.079	-0.021	-0.061	0.109	0.006	0.002	-0.021	-0.014	-0.028	-0.395	0.333	1.000	0.595	0.279
significance (2-sided)		0.226	0.432	0.049	0.147	0.702	0.263	0.044	0.913	0.978	0.698	0.799	0.602	0.000	0.000	.	0.000	0.000
Spearman's r	**dist.rook**	-0.006	-0.073	0.058	-0.057	0.078	0.066	0.009	-0.082	0.082	-0.048	-0.036	0.029	-0.226	0.434	0.595	1.000	0.081
significance (2-sided)		0.914	0.178	0.280	0.289	0.152	0.225	0.866	0.130	0.131	0.380	0.506	0.587	0.000	0.000	0.000	.	0.133
Spearman's r	**dist.sett**	-0.035	-0.012	0.005	0.012	0.142	0.158	-0.017	-0.165	0.168	0.016	0.050	-0.092	-0.197	-0.221	0.279	0.081	1.000
significance (2-sided)		0.519	0.819	0.925	0.831	0.008	0.003	0.755	0.002	0.002	0.765	0.357	0.089	0.000	0.000	0.000	0.133	.

App. A-3. Ranges of the measured values for all presence/absence variables (landscape scale).

Right bars represent nest sites (presence), left bars random points (absence). Tested for differences in distribution with Mann-Whitney U test, Significance levels: p>0.001***, p>0.01**, p>0.05*.

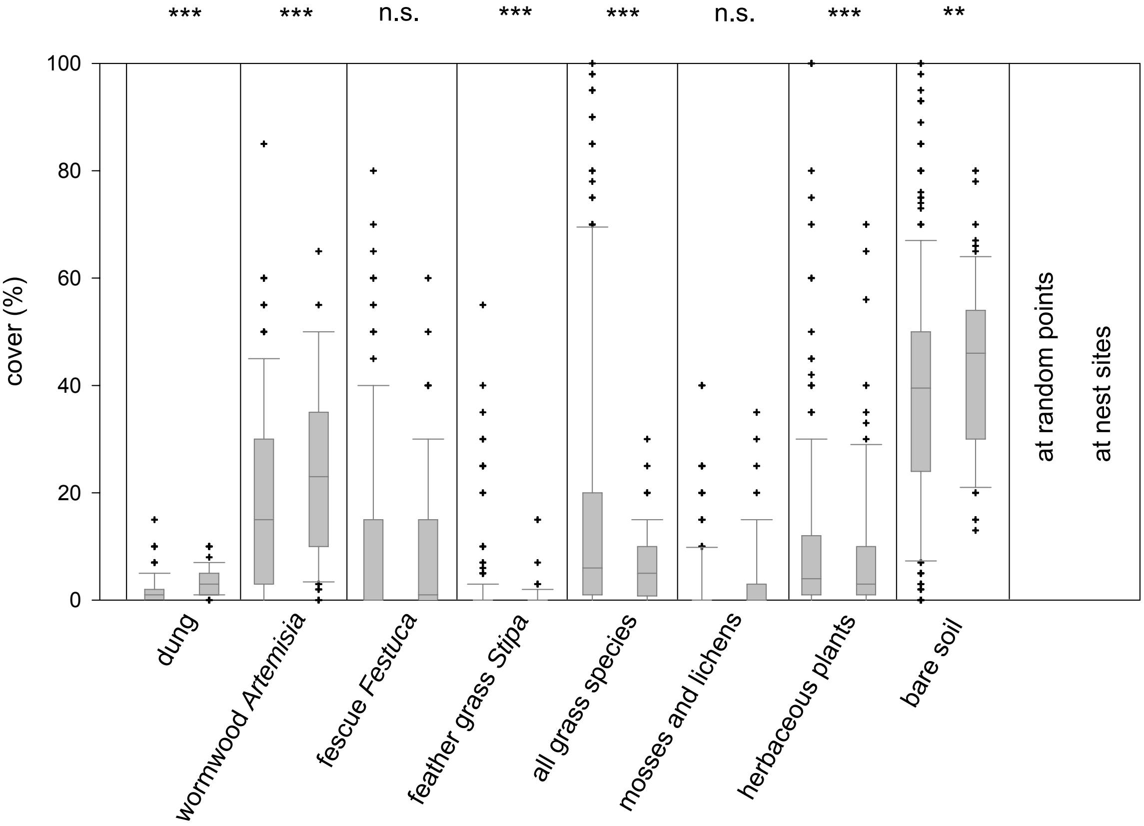


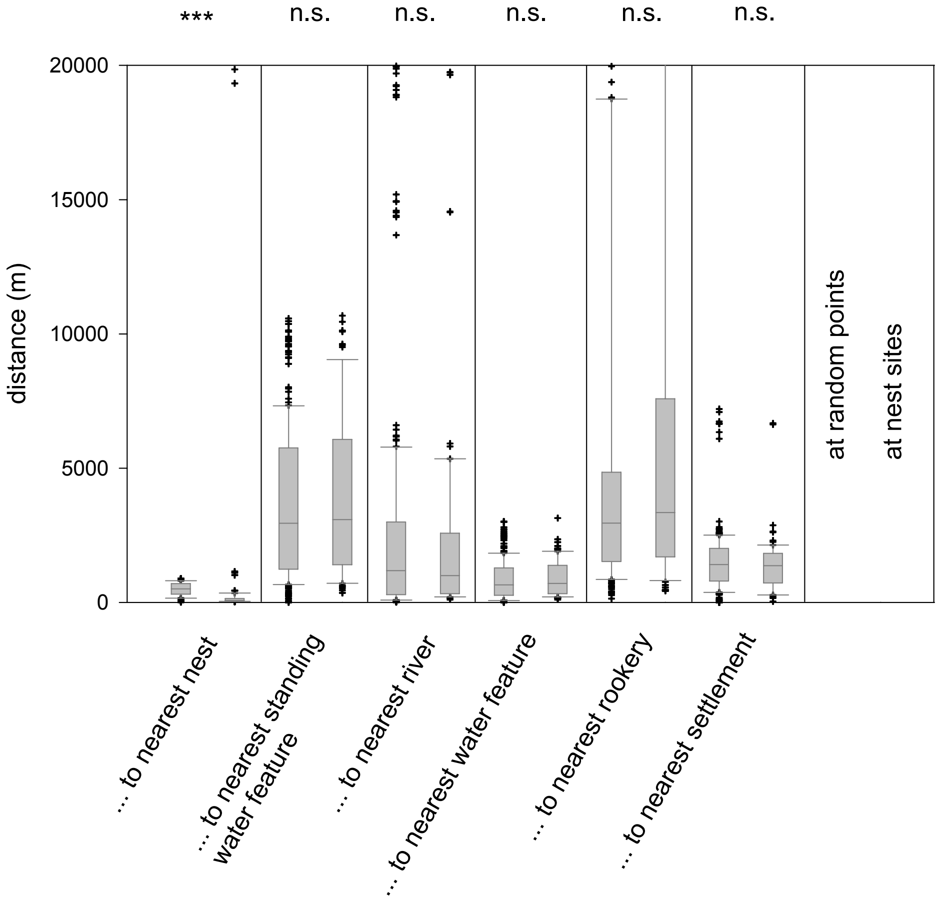

**
n.s.
***
***
***
***
40000
30000
20000
10000
0
distance (m)
at random points
at nest sites
... to nearest colony
... to nearest standing water feature
... to nearest river
... to nearest water feature
... to nearest rookery
... to nearest settlement

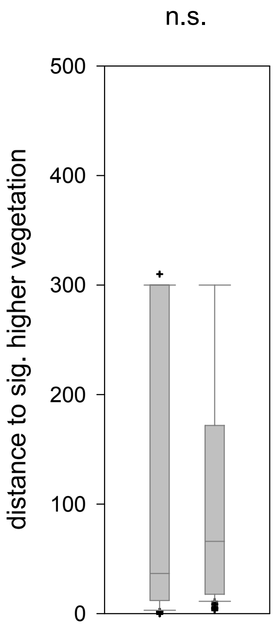

***
***
median
maximum
300
200
100
0
vegetation height (mm)

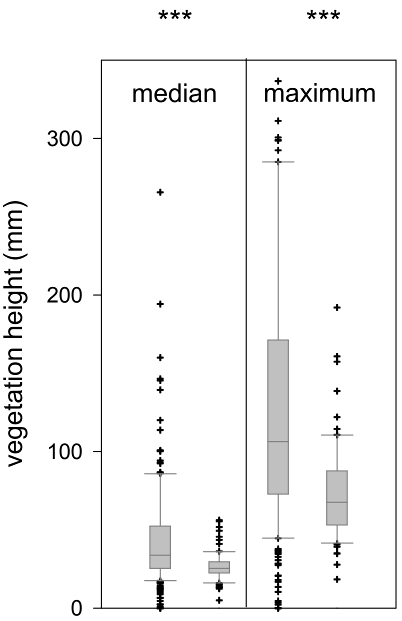
App. A-4. Ranges of the measured values for all presence/absence variables (colony scale).

Right bars represent nest sites (presence), left bars random points (absence). Tested for differences in distribution with Mann-Whitney U test, Significance levels: p>0.001***, p>0.01**, p>0.05*.




***
n.s.
n.s.
n.s.
n.s.
n.s.
distance (m)
20000
15000
10000
5000
0
at random points
at nest sites
… to nearest nest
… to nearest standing water feature
… to nearest river
… to nearest water feature
… to nearest rookery
… to nearest settlement


n.s.
distance to sig. higher vegetation
500
400
300
200
100
0


***
***
median
maximum
vegetation height (mm)
300
200
100
0

App. A-5. Overview of all models considered as “adequate”, landscape scale (cf. chapter 2.3.1)

Shortcuts: AUC = Area under the Receiver-Operator Characteristic, R2N = Nagelkerke’s R², p_Kappa = value, at which Cohen’s κ is maximized, AICc = corrected Akaike’s Information Criterion. “Significance reduction” refers to a Likelihood-ratio test applied to every single model to test if it was better than (or just as good as) any model with one variable less. Models sorted according to increasing AIC_C value.

Model ID	AUC	AUC_lower CI	AUC_upper CI	R2N	Cohen's K	p_Kappa	AICc	Formula	significance reduction
model322	0.99	0.96	1.01	0.89	0.96	0.28	34.30	vegH.mod +dist.riv +dist.sett	0.000
model48	1.00	0.99	1.00	0.91	0.94	0.40	36.20	cov.art +cov.art^2 +cov.gras.tot +cov.gras.tot^2 +dist.wat +dung.tot	0.000
model75	0.99	0.98	1.00	0.88	0.91	0.55	40.10	cov.art +cov.art^2 +vegH.mod +dist.riv +dist.wat	0.000
model207	0.99	0.97	1.00	0.87	0.91	0.61	40.80	vegH.mod +dist.riv +dist.wat +dung.tot	0.000
model326	0.98	0.95	1.01	0.84	0.92	0.37	44.00	vegH.mod +dist.sett +dung.tot	0.066
model270	0.98	0.96	1.00	0.84	0.90	0.30	44.10	cov.stip +dist.riv +dist.sett	0.000
model368	0.97	0.94	1.01	0.82	0.92	0.32	45.30	vegH.mod +dist.sett	0.000
model27	0.99	0.98	1.00	0.85	0.88	0.45	46.00	cov.art +cov.art^2 +cov.stip +dist.wat +dung.tot	0.000
model82	0.99	0.97	1.00	0.85	0.87	0.51	46.00	cov.art +cov.art^2 +dist.riv +dist.wat +dung.tot	0.000
model323	0.98	0.96	1.00	0.83	0.91	0.65	46.10	vegH.mod +dist.riv +dung.tot	0.000
model371	0.98	0.96	1.00	0.81	0.89	0.43	47.70	dist.riv +dist.sett	0.000
model84	0.99	0.97	1.00	0.84	0.87	0.52	47.80	cov.art +cov.art^2 +dist.wat +dist.sett +dung.tot	0.006
model245	0.98	0.97	1.00	0.82	0.88	0.40	49.30	cov.art +cov.art^2 +dist.wat +dung.tot	0.000
model330	0.98	0.95	1.00	0.80	0.88	0.35	51.30	dist.wat +dist.sett +dung.tot	0.060
model319	0.98	0.96	1.00	0.81	0.85	0.49	51.40	cov.veg +cov.veg^2 +dist.wat +dung.tot	0.000
model314	0.97	0.94	1.01	0.81	0.85	0.40	52.50	cov.veg +cov.veg^2 +vegH.mod +dung.tot	0.000
model373	0.97	0.94	1.00	0.78	0.86	0.37	52.70	dist.wat +dist.sett	0.000
model321	0.97	0.95	1.00	0.77	0.86	0.47	56.80	vegH.mod +dist.riv +dist.wat	0.012
model265	0.97	0.94	1.00	0.77	0.88	0.52	57.40	cov.stip +vegH.mod +dist.riv	0.000
model25	0.98	0.95	1.00	0.79	0.82	0.31	58.10	cov.art +cov.art^2 +cov.stip +dist.riv +dung.tot	0.000
model273	0.97	0.95	1.00	0.76	0.81	0.23	58.80	cov.stip +dist.wat +dung.tot	0.000
model300	0.96	0.93	1.00	0.77	0.82	0.49	59.90	cov.herb +cov.veg +cov.veg^2 +dung.tot	0.000
model366	0.96	0.92	1.00	0.74	0.86	0.52	61.00	vegH.mod +dist.riv	0.000
model365	0.95	0.91	0.99	0.75	0.80	0.45	61.70	cov.veg +cov.veg^2 +dung.tot	0.000
model243	0.97	0.94	1.00	0.76	0.80	0.54	61.90	cov.art +cov.art^2 +dist.riv +dung.tot	0.000
model13	0.97	0.95	1.00	0.76	0.78	0.31	62.90	cov.art +cov.art^2 +cov.stip +cov.herb +dung.tot	0.000
model218	0.97	0.94	0.99	0.74	0.78	0.54	64.30	cov.art +cov.art^2 +cov.stip +dung.tot	0.000
model259	0.97	0.94	0.99	0.73	0.78	0.45	64.50	cov.stip +cov.herb +dung.tot	0.000
model271	0.96	0.92	1.00	0.73	0.80	0.31	64.90	cov.stip +dist.riv +dung.tot	0.000
model309	0.96	0.94	0.99	0.72	0.75	0.36	65.50	cov.herb +dist.wat +dung.tot	0.000
model216	0.96	0.93	0.99	0.72	0.76	0.38	67.80	cov.art +cov.art^2 +cov.stip +dist.wat	0.000
model328	0.96	0.93	0.99	0.71	0.80	0.28	67.80	dist.riv +dist.wat +dung.tot	0.000
model231	0.96	0.93	0.99	0.72	0.78	0.58	68.20	cov.art +cov.art^2 +cov.herb +dung.tot	0.000
model241	0.96	0.93	0.99	0.72	0.76	0.36	68.30	cov.art +cov.art^2 +dist.riv +dist.wat	0.000
model374	0.96	0.93	0.99	0.69	0.73	0.23	69.10	dist.wat +dung.tot	0.000
model339	0.95	0.92	0.99	0.70	0.74	0.37	69.40	cov.art +cov.art^2 +dung.tot	0.000
model307	0.95	0.90	1.00	0.70	0.79	0.39	70.10	cov.herb +dist.riv +dung.tot	0.000
model347	0.96	0.93	0.99	0.68	0.74	0.41	70.20	cov.stip +dung.tot	0.000

model337	0.94	0.90	0.98	0.66	0.70	0.50	75.90	cov.art +cov.art^2 +dist.wat	0.000
model360	0.94	0.91	0.98	0.65	0.71	0.37	76.00	cov.herb +dung.tot	0.000
model115	0.95	0.91	0.99	0.67	0.76	0.47	77.00	cov.stip +cov.herb +dist.riv +dist.wat	0.002
model367	0.94	0.90	0.98	0.64	0.77	0.46	77.10	vegH.mod +dist.wat	0.000
model372	0.94	0.88	0.99	0.63	0.74	0.21	77.60	dist.riv +dung.tot	0.000
model269	0.93	0.89	0.97	0.61	0.69	0.34	83.40	cov.stip +dist.riv +dist.wat	0.001
model256	0.93	0.89	0.98	0.61	0.69	0.37	84.00	cov.stip +cov.herb +dist.riv	0.000
model305	0.92	0.88	0.97	0.58	0.68	0.60	88.00	cov.herb +dist.riv +dist.wat	0.000
model257	0.91	0.86	0.96	0.57	0.67	0.34	89.00	cov.stip +cov.herb +dist.wat	0.000
model255	0.92	0.87	0.98	0.56	0.79	0.28	91.00	cov.stip +cov.herb +vegH.mod	0.000
model344	0.91	0.85	0.97	0.53	0.68	0.47	92.10	cov.stip +dist.riv	0.000
model336	0.91	0.85	0.97	0.55	0.66	0.53	92.60	cov.art +cov.art^2 +dist.riv	0.000
model370	0.90	0.84	0.95	0.51	0.59	0.55	95.40	dist.riv +dist.wat	0.000
model356	0.91	0.85	0.97	0.51	0.72	0.27	95.70	cov.herb +vegH.mod	0.000
model254	0.90	0.85	0.95	0.54	0.62	0.33	96.00	cov.stip +cov.herb +cov.veg +cov.veg^2	0.000
model345	0.88	0.82	0.93	0.50	0.59	0.31	96.70	cov.stip +dist.wat	0.000
model343	0.90	0.85	0.96	0.50	0.67	0.29	97.40	cov.stip +vegH.mod	0.000
model249	0.91	0.85	0.96	0.53	0.64	0.37	97.50	cov.stip +cov.gras.tot +cov.gras.tot^2 +vegH.mod	0.000
model357	0.90	0.83	0.97	0.49	0.62	0.43	98.40	cov.herb +dist.riv	0.000
model275	0.91	0.86	0.96	0.52	0.65	0.34	100.00	cov.gras.tot +cov.gras.tot^2 +cov.herb +cov.veg +cov.veg^2	0.001
model358	0.87	0.81	0.93	0.47	0.61	0.43	100.60	cov.herb +dist.wat	0.000
model212	0.88	0.82	0.94	0.47	0.56	0.40	105.20	cov.art +cov.art^2 +cov.stip +cov.herb	0.002
model342	0.85	0.78	0.91	0.42	0.50	0.31	109.70	cov.stip +cov.veg +cov.veg^2	0.000
model331	0.84	0.78	0.91	0.39	0.47	0.42	112.70	cov.art +cov.art^2 +cov.stip	0.000
model355	0.84	0.77	0.91	0.38	0.47	0.40	114.40	cov.herb +cov.veg +cov.veg^2	0.000
model333	0.83	0.75	0.91	0.35	0.49	0.35	117.40	cov.art +cov.art^2 +cov.herb	0.001

App. A-6. Overview of all models considered as “adequate”, colony scale (cf. chapter 2.3.1)

Shortcuts: AUC = Area under the Receiver-Operator Characteristic, R2N = Nagelkerke’s R², p_Kappa = value, at which Cohen’s κ is maximized, AICc = corrected Akaike’s Information Criterion. “Significance reduction” refers to a Likelihood-ratio test applied to every single model to test if it was better than (or just as good as) any model with one variable less. Models sorted according to increasing AIC$_C$ value.

Model ID	AUC	AUC_lower CI	AUC_upper CI	R2N	Cohen's K	p_Kappa	AICc	Formula	significance reduction
model4	0.92	0.87	0.96	0.56	0.7	0.41	218.1	cov.dung +cov.dung^2 +cov.veg +cov.veg^2 +vegH.max +dist.nest	0.000
model7	0.81	0.76	0.86	0.31	0.42	0.39	296.8	cov.dung +cov.dung^2 +cov.grass +vegH.max	0.000
model9	0.82	0.77	0.87	0.34	0.45	0.34	289.8	cov.dung +cov.dung^2 +cov.veg +cov.veg^2 +vegH.max	0.000
model10	0.92	0.87	0.96	0.57	0.7	0.4	213.7	cov.dung +cov.dung^2 +cov.veg +cov.veg^2 +dist.nest	0.000
model22	0.9	0.85	0.95	0.49	0.66	0.42	234.7	cov.grass +dist.nest	0.000
model25	0.91	0.86	0.95	0.52	0.71	0.51	225.9	vegH.max +dist.nest	0.000
model 99	0.89			0.46				dist.nest	

Appendix B. Scripts used for automated model calculations in S-Plus

App. B-1. S-Plus Script: multi-model calculation

A script for S-Plus 6.1, developed by Dr. Barbara Strauss (Landscape Ecology Group, University of Oldenburg. Contact: barbara.strauss@uni-oldenburg.de

This script calculates multivariate logistic regression models for all possible combinations of 4,3, and 2 variables.

```
#Rechnet Modelle für alle möglichen Kombinationen von 4, 3 und 2 Variablen.
#Als Gütemaße werden AUC (incl. Konf.intervall), Nagelkerke's R^2, Kappa und AICc ausgegeben.
#Die Textdatei nach Excel importieren (Komma-getrennt), dort das zugehörige Makro laufen lassen (mit
#Cursor auf der vorletzten Zeile, 1. Spalte)

daten <- landscape                                  # data.frame, der die Daten enthält
art <- "kiebitz.landscape"                          # Spalte mit der zu modellierenden Art
dateiname.out <- c(paste("kombi_",art,".xls",sep="")) #Name für die Output-Datei

Variablen.Datei <- paste(c("univar_",art,".xls"),collapse="") #Name der Datei, die eingelesen wird

korr.max<-0.7 #maximale Korrelation, die zwischen den Variablen eines Modells erlaubt ist


variablen.tabelle <- na.exclude(importData(Variablen.Datei, keep=c(1,2),startRow=1))
                        #einzulesende Spalten aus der einzulesenden Datei
                        #1. einzulesende Spalte enthält die Variabelnnamen
                        #2. einzulesende Spalte enthält die Info, ob die Variable sigmoidal (Wert in Spalte = 1)
                        #oder unimodal (Wert = 2) reagiert

max.p.coeff <- 0.15 #maximaler p-Wert der Modellkoeffizienten
max.p.coeff.pl <- 0.1
max.p.rm <- 0.1         #maximaler p-Wert, für den das Modell als besser betrachtet wird als ein Modell,
                                        #das eine Variable weniger enthält
min.r2n <- 0.3

#unterhalb von hier nichts verändern!
#############################################################################################################
variablen <- as.vector(variablen.tabelle[,1])
var.qu <- na.exclude(as.data.frame(ifelse(variablen.tabelle[,2]==2,as.vector(variablen.tabelle[,1]),NA)))
var.qu <- as.vector(var.qu[,1])

faelle<-nrow(daten)
dummy.1<-rnorm(faelle, mean=10, sd=5) #dummy-Variablen einfügen, falls nur 3 oder 4 Variablen signifikante sind
dummy.2<-rnorm(faelle, mean=40, sd=15)#(das Skript funktioniert erst ab 5 Variablen); aus Output-Date wieder rauslöschen!
dummy.3<-rnorm(faelle, mean=50, sd=6)#(das Skript funktioniert erst ab 5 Variablen); aus Output-Date wieder rauslöschen!
if(length(variablen)==4) {
        daten<-data.frame(daten,dummy.1)
        variablen[5]<-"dummy.1"}
if(length(variablen)==3) {
        daten<-data.frame(daten,dummy.1,dummy.2)
        variablen[4:5]<-c("dummy.1","dummy.2")}
if(length(variablen)==2) {
        daten<-data.frame(daten,dummy.1,dummy.2,dummy.3)
        variablen[3:5]<-c("dummy.1","dummy.2","dummy.3")}

library(logistf, T)
library(Hmisc, T)
library(Design, T)
################################################################################
#eigentlich überflüssig, wird nur benötigt, falls Variablen von Hand eingegeben werden sollen.
#Erklaerende Variablen. Es duerfen keine Luecken ("") sein, außer am Ende.
#variablen <- c(
#       "d.gew","ks.pr","pr.90.hi","streu","off.pr","kfp","lk","kak.eff","ph","brache.ges.100","BGd.hoch.75"
#       )
#var.qu <- c("ks.pr","pr.90.hi","streu")                #Variablen, die auch quadriert beruecksichtigt werden sollen
###############################################################################################
n <- length(variablen)                  #Anzahl unabhängiger Variablen
k <- 4                                              #Max. Anzahl von Parametern im Modell. Kann nicht verändert werden (jedenfalls nicht ohne
                                                        #größere Änderungen im restlichen Skript)
var.qu.length <- length(var.qu)             #Anzahl Variablen mit Quadrat
########################################################################
#Erzeugen von allen Kombinationen von n (=alle erklärenden Variablen)
#mit k Elementen (also alle möglichen Kombis mit k Variablen)
subsets<-function(n, k, set = 1:n){
        if(k <= 0) NULL
        else if(k >= n) set
        else rbind(cbind(set[1], Recall(n-1, k-1, set[-1])), Recall(n-1, k, set[-1]))
}

#Matrizen mit allen Kombinationen von k Variablen
set.4.variablen <- subsets (n, k)
set.3.variablen <- subsets (n, k-1)
set.2.variablen <- subsets (n, k-2)

#Zahlen in Matrix durch Variablenname ersetzen
set.4.variablen.namen<- t(apply (set.4.variablen, 1, function (x, variablen) variablen[x],variablen))
set.3.variablen.namen<- t(apply (set.3.variablen, 1, function (x, variablen) variablen[x],variablen))
```

```
set.2.variablen.namen<- t(apply (set.2.variablen, 1, function (x, variablen) variablen[x],variablen))

#Funktion, um zu berechen, ob Korrelationen zwischen Variablen jedes Modells größer ist als korr.max
korrelationen<-function(varset){
        korreliert<-vector(mode="numeric", length=nrow(varset))
        for(i in 1: nrow(varset)){
                zeile<-varset[i,]
                zu.korrelieren<-as.matrix(daten[zeile])
                korrelationen<-rcorr(zu.korrelieren,type="spearman")
                ohne.1<-as.vector(ifelse(korrelationen$r==1,0,korrelationen$r))
                if(max(ohne.1)>korr.max)  korreliert[i]<-1
        }
        varset<-cbind(varset,korreliert)
}

#Funktion zum Korrelationen markieren anwenden
set.4.variablen.korrelationen<-korrelationen(set.4.variablen.namen)
set.3.variablen.korrelationen<-korrelationen(set.3.variablen.namen)
set.2.variablen.korrelationen<-korrelationen(set.2.variablen.namen)

modelle.mit.korrelationen<- c(set.4.variablen.korrelationen[,5],
        set.3.variablen.korrelationen[,4], set.2.variablen.korrelationen[,3])

#Liste mit den Variablennamen erzeugen. Diese wird später gebraucht, um für die Daten jedes Modells
#einen Data.Frame erzeugen zu können, aus dem die NAs ausgeschlossen werden (sonst funktioniert logistf nicht)
set.namen <- matrix(data=NA,ncol=4, nrow=nrow(set.4.variablen.namen)+nrow(set.3.variablen.namen)+nrow(set.2.variablen.namen))
combis.4var<-nrow(set.4.variablen.namen)
combis.3var<-nrow(set.3.variablen.namen)
combis.2var<-nrow(set.2.variablen.namen)
set.namen[1:combis.4var,1:4]<-set.4.variablen.namen[,1:4]
set.namen[(combis.4var+1):(combis.4var+combis.3var),1:3]<-set.3.variablen.namen[,1:3]
set.namen[(combis.4var+combis.3var+1):nrow(set.namen),1:2]<-set.2.variablen.namen[,1:2]

#Funktion, mit der fuer die unimodalen Variablen auch die quadrierten
#Terme eingefuegt werden
quadrate.einfuegen <- function(varset)
        {       col.count <- ncol(data.frame(varset))
                row.count <- nrow(data.frame(varset))
                for (i in 1:var.qu.length)
                {       X <- var.qu[i]
                        Xneu <-paste(X, " +",X,"^2", sep="")
                        for (col in 1:col.count)                                {       for (row in 1:row.count)
                                {       if (varset[row,col] == X) varset[row,col] <- Xneu
                                }
                        }
                }
                varset
        }

#quadrate.einfuegen auf die Variablen-Sets anwenden
if(length(var.qu)>0) varset.4.variablen.matrix<-quadrate.einfuegen(set.4.variablen.namen[,1:4])
        if(length(var.qu)==0) varset.4.variablen.matrix<-set.4.variablen.namen[,1:4]
if(length(var.qu)>0) varset.3.variablen.matrix<-quadrate.einfuegen(set.3.variablen.namen[,1:3])
        if(length(var.qu)==0) varset.3.variablen.matrix<-set.3.variablen.namen[,1:3]
if(length(var.qu)>0) varset.2.variablen.matrix<-quadrate.einfuegen(set.2.variablen.namen[,1:2])
        if(length(var.qu)==0) varset.2.variablen.matrix<-set.2.variablen.namen[,1:2]

#alle Variablen jeweils in einem Ausdruck vereinigen
varset.4.variablen <- apply (varset.4.variablen.matrix, 1, paste, collapse = " +")
varset.3.variablen <- apply (varset.3.variablen.matrix, 1, paste, collapse = " +")
varset.2.variablen <- apply (varset.2.variablen.matrix, 1, paste, collapse = " +")


#Ausdrücke von oben, aber mit einer Variable weniger
        varset.4.variablen.minus1 <- paste("resp ~ ", apply(varset.4.variablen.matrix[,2:4], 1, paste, collapse = " +"))
        varset.4.variablen.minus2 <- paste("resp ~ ", apply(varset.4.variablen.matrix[,c(1,3,4)], 1, paste,collapse = " +"))
        varset.4.variablen.minus3 <- paste("resp ~ ", apply(varset.4.variablen.matrix[,c(1,2,4)], 1, paste,collapse = " +"))
        varset.4.variablen.minus4 <- paste("resp ~ ", apply(varset.4.variablen.matrix[,1:3], 1, paste,collapse = " +"))
        varset.4.variablen.minus <- data.frame(varset.4.variablen.minus1,varset.4.variablen.minus2,
                                                                        varset.4.variablen.minus3,varset.4.variablen.minus4)
        varset.4.variablen.minus <- as.matrix(varset.4.variablen.minus)

        varset.3.variablen.minus1 <- paste("resp ~ ", apply (varset.3.variablen.matrix[,c(2,3)], 1, paste, collapse = " +"))
        varset.3.variablen.minus2 <- paste("resp ~ ", apply (varset.3.variablen.matrix[,c(1,3)], 1, paste, collapse = " +"))
        varset.3.variablen.minus3 <- paste("resp ~ ", apply (varset.3.variablen.matrix[,c(1,2)], 1, paste, collapse = " +"))
        varset.3.variablen.minus <- data.frame(varset.3.variablen.minus1, varset.3.variablen.minus2 , varset.3.variablen.minus3)
        varset.3.variablen.minus[,4] <- NA
        varset.3.variablen.minus <- as.matrix(varset.3.variablen.minus)

        varset.2.variablen.minus1 <- paste("resp ~ ", varset.2.variablen.matrix[,2])
        varset.2.variablen.minus2 <- paste("resp ~ ", varset.2.variablen.matrix[,1])
        varset.2.variablen.minus <- data.frame(varset.2.variablen.minus1,varset.2.variablen.minus2)
        varset.2.variablen.minus[,3:4] <- NA
        varset.2.variablen.minus <- as.matrix(varset.2.variablen.minus)


form.minus <- rbind(varset.4.variablen.minus,varset.3.variablen.minus,varset.2.variablen.minus)

#Formeln für alle zu rechnenden Modelle
form <- c(paste("resp ~ ", varset.4.variablen),
        paste("resp ~ ", varset.3.variablen), paste("resp ~ ", varset.2.variablen))

#Formeln ohne "resp ~" für Ausgabe in Datei
form.out <- c(varset.4.variablen, varset.3.variablen, varset.2.variablen)

######################################################################################
#"Ankreuzliste", welche Variablen in jedem Modell verwendet wurden

#Tabelle mit 1 Reihe für jede Variablenkombination und so vielen Spalten,
#wie Variablen. Wenn die Variable in der entsprechenden Kombi vorkommt,
#steht in dem Feld eine Zahl
form.tab0 <- matrix (nrow = nrow(set.4.variablen.namen), ncol = n)
dimnames(form.tab0)[[2]] <- c(variablen)
```

```
form.tab1 <- matrix (nrow = nrow(set.3.variablen.namen), ncol = n)
form.tab2 <- matrix (nrow = nrow(set.2.variablen.namen), ncol = n)

var.in.spalten <- function(form.tab,t.set){
        for (col in 1:ncol(form.tab)){
                print(col)
                v <- variablen[col]
                        for (row in 1:nrow(form.tab)){
                                form.tab[row,col] <- charmatch(v, c(t.set[row,]), nomatch="")
                        }
        }
        form.tab
}

form.tab0 <- var.in.spalten(form.tab0,set.4.variablen.namen[,1:4])
form.tab1 <- var.in.spalten(form.tab1,set.3.variablen.namen[,1:3])
form.tab2 <- var.in.spalten(form.tab2,set.2.variablen.namen[,1:2])
form.tab <- data.frame(rbind(form.tab0, form.tab1, form.tab2))
##############################################################
#Data.Frames für Output anlegen, Modelle berechnen und Ergebnisse exportieren

#Anzahl Zeilen (d.h. Anzahl möglicher Kombinationen )
combi.count<-length(form)

#Funktion zur Berechnung der Modellgüte
guete.logistf <- function(Modell){

        #zur Berechnung notwendige Infos aus dem Modell extrahieren
        ###################
        y.pred <- Modell$predict
        y.obs <- Modell$y
        n<-Modell$n


        #AUC
        ##############    auc.statistik<-rcorr.cens(y.pred,y.obs,outx=FALSE)
        auc<-auc.statistik[1]
        auc.u<-auc.statistik[1]-auc.statistik[3]
        auc.o<-auc.statistik[1]+auc.statistik[3]


        #P-value Modell
        ###########################
        #1.) loglikelihood Modell
        ll.modell <- sum(na.exclude(y.obs*logb(y.pred)+((1-y.obs)*logb(1-y.pred))))

        #2.) loglikelihood Nullmodell (nur mit Konstante)
        eins <- sum(y.obs) #Summe der Vorkommen
        prevalence <- eins/length(y.obs) #Prävalenz
        y.pred.0 <- prevalence
        ll.null <- sum(na.exclude(y.obs*logb(y.pred.0)+((1-y.obs)*logb(1-y.pred.0))))

        #3.) aus den Likelihoods die Devianz berechnen
        devianz.null<- ll.null*(-2)
        devianz.modell<- ll.modell*(-2)

        #4.) Likelihood Ratio und daraus Signifkanz berechnen (Chiquadrat-Test)
        LR <- devianz.null - devianz.modell #Likelihood Ratio
        pval <- 1-pchisq(LR, Modell$df)

        #AICc
        ##################
        p<-length(Modell$coefficients)
        aic<- (-2*ll.modell) + (2*p)
        aicc<- aic + ((2*p*(p+1))/(n-p-1))

        #R2N
        ######################################
        l.null <- exp(ll.null)
        l.modell <-exp(ll.modell)
        RCox <- 1-((l.null/l.modell)^(2/n))
        RMax <- 1-(l.null^(2/n))
        r2n <- RCox/RMax

        #Kappa
        ###########################################
        #1.) Funktion zur Berechnung von Kappa bei einem gegebenen Schwellenwert
        kappa.eval<-function(y.obs,y.pred,i){
        temp.vector<-y.obs+(as.numeric((y.pred>=i))*10)
        cont.a<-sum(temp.vector==11)
        cont.b<-sum(temp.vector==10)
        cont.c<-sum(temp.vector==1)
        cont.d<-sum(temp.vector==0)
        n<-cont.a+cont.b+cont.c+cont.d
        p.obs<-cont.a/n+cont.d/n
        p.exp<-((cont.a+cont.c)*(cont.a+cont.b)+(cont.c+cont.d)*(cont.b+cont.d))/n^2
        kappa.value<-(p.obs-p.exp)/(1-p.exp)
        return(kappa.value)
        }

        #2.) Kappa für alle Schwellenwerte (pcrit) berechenen und den Schwellenwert mit
        #maximalem Kappa bestimmen
         thresholds<-seq(0,1,0.01)
         pcrit<-vector(length=length(thresholds))
        for (j in 1:length(thresholds)){
                 pcrit[j]<-kappa.eval(y.obs,y.pred,thresholds[j])
        }
    kappa.max<-max(pcrit)
    p.kappa<-min(thresholds[pcrit>=max(pcrit)])

        #Ergebnis, das von der Güte-Funktion zurückgegeben wird
        data.frame(auc,auc.u,auc.o,r2n,aicc,kappa.max,p.kappa)
}
```

```
#matrices for output - later put in data frame
out <- matrix (nrow = combi.count, ncol = 10, dimnames=list(paste("model",1:combi.count,sep="")))
dimnames(out)[[2]]<- c("AUC","AUC.u","AUC.o","R2N","AICc","Kappa","PKappa","Formula","sig.weniger","sig.coeffs")
out<-data.frame(out)

#für alle Kombinationen die Modelle und deren Güte berechnen und das Ergebnis in out ablegen
#i<-24 rm<-1 a<-3
for (i in 1:combi.count){
        if(modelle.mit.korrelationen[i]==0) {
                #1. Vektor mit den Namen der erklärenden Variablen erzeugen (ohne "NA")
                        indep<-vector(mode="character")
                                for (a in 1:4){
                                        if (set.namen[i,a] != "NA")  indep[a]=set.namen[i,a]
                                }
                #2. Data.Frame mit allen für das Modell benötigten Daten anlegen
                        data<-data.frame(daten[,art])
                        data<-na.exclude(data.frame(data,daten[,indep]))
                        resp<-data[,1]
                        volles.modell <- logistf(formula=form[i]    ,data=data,pl=F,firth=T)
                        devianz.volles.modell <- -2*(sum(na.exclude(volles.modell$y*logb(volles.modell$predict)+
                                                                                ((1-volles.modell$y)*logb(1-volles.modell$predict)))))
                        out[i,1:7] <- c(guete.logistf(volles.modell))
                        out[i,10]<-max(volles.modell$prob)

                        for (rm in 1:4){
                                if(out[i,4]<min.r2n) out[i,10]<-1
                                if(out[i,10]>max.p.coeff) out[i,10]<-1
                                if(out[i,10]==1) break
                                if (form.minus[i,rm]!="NA") {
                                        reduced.model<-logistf(formula=form.minus[i,rm], data=data,pl=F,firth=T)
                                        devianz.reduced.model <- -2*(sum(na.exclude(reduced.model$y*logb(reduced.model$predict)+
                                                                                                        ((1-reduced.model$y)*logb(1-reduced.model$predict)))))
                                        LR <- devianz.reduced.model - devianz.volles.modell #Likelihood Ratio
                                        df <- volles.modell$df - reduced.model$df
                                        pval <- 1-pchisq(LR, df)
                                        out[i,9]<-pval
                                        if(pval>max.p.rm) out[i,9]<-1
                                        #if(out[i,9]==1) break
                                        #out[i,10]<-max(logistf(formula=form[i],data=data,pl=T,firth=T,maxstep=0.5)$prob)
                                        #if(out[i,10]>max.p.coeff.pl) out[i,10]==1
                                }
                        }
                        out[i,8] <- form.out[i]
                        print(paste(i, "von", combi.count, collapse = " "))
        }
}

#Output in out.final kombinieren und Zahlen runden
out.alles <- data.frame(round(out[c(1,2,3,4,6,7)], digits=2),round(out[5], digits=1),out[8],out[9],out[10],form.tab)

out.final<-out.alles[1,]
        for (i in 1:nrow(out.alles)){
                        ausschliessen<-max(na.exclude(as.data.frame(as.numeric(c(out.alles[i,9],out.alles[i,10],modelle.mit.korrelationen[i])))))
                        if(ausschliessen!=1)
                out.final<-rbind(out.final,out.alles[i,])
        }
out.final<-out.final[-1,]
#exportiert die Tabelle im Excel-Format, wird im Default-Ordner abgelegt
exportData(out.final, dateiname.out, type="EXCEL", rowNames=T, colNames=T)
```

App. B-2. S-Plus Script: model averaging

A script for S-Plus 6.1, developed by Dr. Barbara Strauss (Landscape Ecology Group, University of Oldenburg. Contact: barbara.strauss@uni-oldenburg.de

This script can be used for model averaging (cf. BURNHAM & ANDERSON 2002).

SCRIPT 2: Model averaging

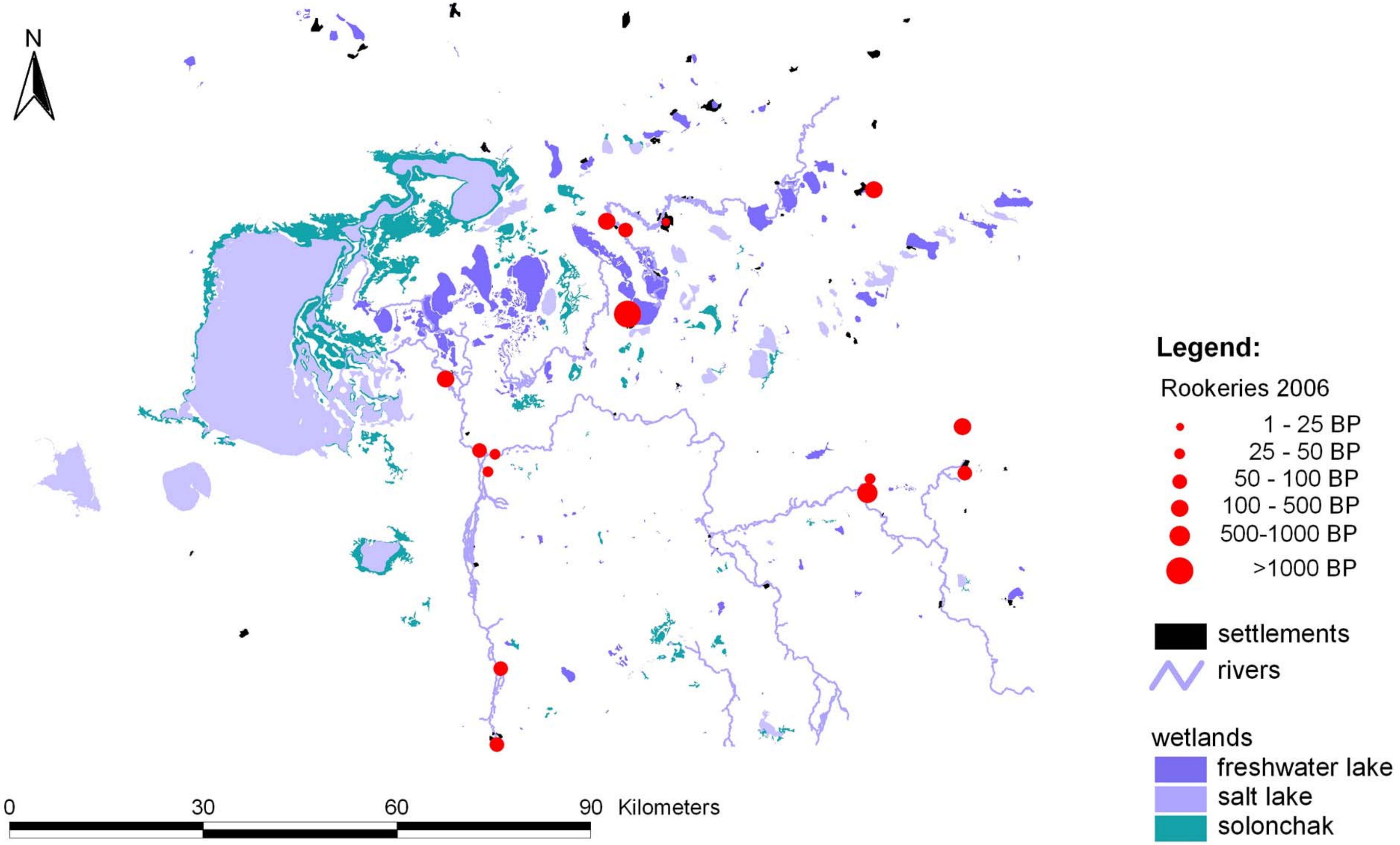
```
#model-averaging, auch, wenn's nur 1 Modell ist (damit der Output der gleiche ist, geaveraged wird da
#natürlich nix.

daten <- colony                                        #data.frame, aus dem die Daten kommen
art <- "kiebitz.colony"

Datei <- paste(c("kombi_",art,".xls"),collapse="") #Name der Excel-Datei, in der die Formeln der einzelenen Modelle stehen
#Datei <- paste(c("berlin_kombi_",art,"_alt.xls"),collapse="") #Name der Excel-Datei, in der die Formeln der einzelenen Modelle stehen

#######################################################
Dateiname.mod.gew <- paste(c("av_gew_",art,".xls"),collapse="") #Name der Excel-Datei, in die AICcs und Gewichte der einzelnen Modelle exportiert werden
Dateiname.koeff <- paste(c("av_koeff_",art,".xls"),collapse="") #Name der Excel-Datei, in die die gemittelten Koeffizienten exportiert werden
Dateiname.guete <- paste(c("av_guete_",art,".xls"),collapse="") #Name der Excel-Datei, in die die Güte des gemittelten Modells exportiert werden

#################################################################
library(Hmisc,T)
library(Design,T)
library(logistf,T)

formeln <- as.matrix(na.exclude(importData(Datei,keep=2,startRow=1))) #Formeln einlesen

var.namen<-as.matrix(na.exclude(importData(Datei,keep=3,startRow=1))) #Variablennamen einlesen
var.namen<-unique(paste(var.namen[,1],collapse=","))
var.namen<-unique(unlist(unpaste(var.namen,sep=",")))



data.art<-na.exclude(data.frame(daten[var.namen],daten[art])) #Datensatz für die Variablen der Art ohne NAs


#Namen der quadrierten Variable, so wie sie nachher von logistf ausgespuckt werden
var.namen.qu<-vector(mode="character", length=length(var.namen))
for (i in 1:length(var.namen)){
        var.namen.qu[i]<-paste(c("I(",var.namen[i],"^2)"),collapse="")
}

werte.einzelmodelle<- matrix(nrow=length(formeln[,1]),ncol=2*length(var.namen)+3) #Matrix für Ergebnisse anlegen
dimnames(werte.einzelmodelle)[[2]]<-c("formel","AICc",var.namen,var.namen.qu,"(Intercept)")

############################################################################
#Schleife über alle Einzelmodelle
for(i in 1: length(formeln)){

        #Modell rechnen
        formel<-paste(art,"~",formeln[i])
        modell <- logistf(formula=formel,data=data.art,pl=F,firth=T)

        #AICc für modell berechnen

        y.pred<-modell$predict
        n<-length(y.pred)
        y.obs<-modell$y
        y.pred.neu<-ifelse(y.pred>0.9999999999999999,0.9999999999999999,y.pred)
        y.pred.neu<-ifelse(y.pred<0.0000000000000000,0.0000000000000000,y.pred.neu)
        ll.modell <- sum(na.exclude(y.obs*logb(y.pred.neu)+((1-y.obs)*logb(1-y.pred.neu))))
        k<-length(modell$coefficients)+1  #Anzahl der Parameter ist Koeffizienten + Konstante + sigma (=Varianzparameter)
                                                                                # siehe Burnham (2003) S. 12
        aic<- (-2*ll.modell) + (2*k)
        aicc<- aic + ((2*k*(k+1))/(n-k-1))   werte.einzelmodelle[i,"AICc"] <- aicc

        #Werte der Koeffizienten in Matrix schreiben
        for (koeff in 1:length(modell$coefficients)){
                werte.einzelmodelle[i,names(modell$coefficients[koeff])]<-modell$coefficients[koeff]
        }
        print(i) #Fortschrittsanzeige

}

modell.nr<-seq(1,length(formeln),1) #Spalte mit Modell-Nr. einfügen, um ursprüngliche Sortierung wiederherstellen zu können
werte.einzelmodelle<-cbind(modell.nr,werte.einzelmodelle)
sortiert.aicc<-sort.col(werte.einzelmodelle,"@ALL",columns.to.sort.by="AICc") #nach AICc sortieren
koeffizienten.werte<-sortiert.aicc[,4:ncol(sortiert.aicc)] #Werte der Koeffizieneten
delta.i <- vector(mode="numeric",length=length(formeln)) #Vektor for delta.i anlegen

#delta.i für jedes Modell berechnen
if(length(formeln)>1){
        for (i in 2:length(formeln)){
                delta.i[i]<-as.numeric(sortiert.aicc[i,"AICc"]) - as.numeric(sortiert.aicc[1,"AICc"])
        }
}
if(length(formeln)==1) delta.i[1]<-0

likelihood<-exp(-0.5*delta.i) #aus delta.i die Likelihood für die gegebenen Daten berechnen (Burnham S. 74)
```

```
wi<-likelihood/sum(likelihood) #und daraus die Akaike Weights (Burnham S. 75)

if(length(formeln)>1) {
        mod.gew<-data.frame(sortiert.aicc[,1:3],delta.i,wi)
        mod.gew<-sort.col(mod.gew,"@ALL",columns.to.sort.by="modell.nr")
        mod.gew<-data.frame(formeln,mod.gew$AICc,mod.gew$delta.i,mod.gew$wi)
}

if(length(formeln)==1) {mod.gew<-c(sortiert.aicc[,1:3],delta.i,wi)
mod.gew<-c(formeln,mod.gew[3:5])
}

names(mod.gew)<-c("formeln","AICc","delta.i","wi")
exportData(mod.gew, Dateiname.mod.gew, type="EXCEL",colNames=T)   #speichert mod.gew
#if(length(formeln)==1){
#       mod.gew<-data.frame(mod.gew)
#       mod.gew<-t(mod.gew)
#       exportData(mod.gew, Dateiname.mod.gew, type="EXCEL",colNames=T)   #speichert mod.gew
#}

########################################################################################
#Koeffizienten mitteln
beta.mittel<-vector(mode="numeric") #Vektor für Mittelwerte der Koeffizienten anlegen (Mitte nur aus Modellen, die die Variable enthalten)
beta.mittel.0<-vector(mode="numeric") #Vektor für Mittelewrte der Koeff anlegen (Mittel aus allen Modellen)
weight.var<-vector(mode="numeric") #Vektor für das Variablengewicht anlegen (Summe der Modellgewichte aller Modelle, die die Variable enthalten)

if(length(formeln)>1){
        for (i in 1:ncol(koeffizienten.werte)){
                koeff.in.modellen<-na.exclude(data.frame(wi,as.numeric(koeffizienten.werte[,i])*wi))
                beta.mittel[i]<-sum(koeff.in.modellen[,2])/sum(koeff.in.modellen[,1])
                beta.mittel.0[i]<-beta.mittel[i]*sum(koeff.in.modellen[,1])
                weight.var[i]<-sum(koeff.in.modellen$wi)
        }
}

if(length(formeln)==1){
        for (i in 1:length(koeffizienten.werte)){
                koeff.in.modellen<-koeffizienten.werte[i]
                beta.mittel[i]<-koeff.in.modellen
                beta.mittel.0[i]<-koeff.in.modellen[1]
                weight.var[i]<-1
        }

}

var.namen.qu.ohneKlammer<-vector(mode="character", length=length(var.namen))
for (i in 1:length(var.namen)){
        var.namen.qu.ohneKlammer[i]<-paste(c(var.namen[i],"^2"),collapse="")
}

var.namen.alle<-c(var.namen,var.namen.qu.ohneKlammer,"Intercept")

out<-data.frame(var.namen.alle,beta.mittel,beta.mittel.0,weight.var)
out<-na.exclude(out)
exportData(out, Dateiname.koeff, type="EXCEL",colNames=T)       #speichert out

#################################################################################
#aus gemittelten Koeffizienten die gemittelte Modellvorhersage und deren Güte berechnen

produkte<-matrix(nrow=nrow(data.art),ncol=nrow(out)-1) #Matrix anlegen für beta*Variablen

#jede Variable mit ihrem Koeffizienten multiplizieren
for (i in 1:ncol(produkte)){
        if(is.element(out[i,1],var.namen)) produkte[,i]<-data.art[,as.character(out[i,1])]*out[i,2]
        if(is.element(out[i,1],var.namen.qu.ohneKlammer))
                        produkte[,i]<-data.art[,unpaste(as.character(out[i,1]),sep="^")[[1]]]^2*out[i,2]
}

summe.produkte<-rowSums(produkte,dims=1)+out[nrow(out),2] #Werte der einzelnen Variablen zusammenzählen
pred1<-1/(1+exp(-1*summe.produkte)) #predicted values berechnen

#das Gleiche für die über alle Modelle gemittelten Koeffizienten
produkte<-matrix(nrow=nrow(data.art),ncol=nrow(out)-1)
for (i in 1:ncol(produkte)){
        if(is.element(out[i,1],var.namen)) produkte[,i]<-data.art[,as.character(out[i,1])]*out[i,3]
        if(is.element(out[i,1],var.namen.qu.ohneKlammer))
                        produkte[,i]<-data.art[,unpaste(as.character(out[i,1]),sep="^")[[1]]]^2*out[i,3]
}
summe.produkte<-rowSums(produkte,dims=1)+out[nrow(out),3]
pred2<-1/(1+exp(-1*summe.produkte))

obs<-data.art[,art]

guete.logistf <- function(y.obs,y.pred){

        n<-length(y.obs)

        #AUC
        ##############   auc.statistik<-rcorr.cens(y.pred,y.obs,outx=FALSE)
        auc<-auc.statistik[1]
        auc.u<-auc.statistik[1]-auc.statistik[3]
        auc.o<-auc.statistik[1]+auc.statistik[3]

        #R2N
        ############################

                y.pred.neu<-ifelse(y.pred>0.9999999999999999,0.9999999999999999,y.pred)
                y.pred.neu<-ifelse(y.pred<0.0000000000000000,0.0000000000000000,y.pred.neu)

                ll.modell <- sum(na.exclude(y.obs*logb(y.pred.neu)+((1-y.obs)*logb(1-y.pred.neu))))
                l.modell<-exp(ll.modell)

                eins <- sum(y.obs) #Summe der Vorkommen             y.pred.0 <- eins/n #Prävalenz
```

```
                    ll.null <- sum(na.exclude(y.obs*logb(y.pred.0)+((1-y.obs)*logb(1-y.pred.0))))
                    l.null<-exp(ll.null)

                    RCox <- 1-((l.null/l.modell)^(2/n))
                    RMax <- 1-(l.null^(2/n))
                    r2n <- RCox/RMax
                    r2n<-ifelse(r2n<0,0,r2n)

          #Kappa
          #########################################
          #1.) Funktion zur Berechnung von Kappa bei einem gegebenen Schwellenwert
          kappa.eval<-function(y.obs,y.pred,i){
          temp.vector<-y.obs+(as.numeric((y.pred>=i))*10)
          cont.a<-sum(temp.vector==11)
          cont.b<-sum(temp.vector==10)
          cont.c<-sum(temp.vector==1)
          cont.d<-sum(temp.vector==0)
          n<-cont.a+cont.b+cont.c+cont.d
          p.obs<-cont.a/n+cont.d/n
          p.exp<-((cont.a+cont.c)*(cont.a+cont.b)+(cont.c+cont.d)*(cont.b+cont.d))/n^2
          kappa.value<-(p.obs-p.exp)/(1-p.exp)
                    sensi<-cont.a/(cont.a+cont.c)
                    spezi<-cont.d/(cont.b+cont.d)
                    pr.korr<-(cont.a+cont.d)/n
          return(kappa.value,sensi,spezi,pr.korr)
          }

          #2.) Kappa für alle Schwellenwerte (pcrit) berechenen und den Schwellenwert mit
          #maximalem Kappa bestimmen

          thresholds<-seq(0,1,0.01)
          pcrit<-matrix(ncol=4,nrow=length(thresholds))
          for (j in 1:length(thresholds)){
       pcrit[j,]<-kappa.eval(y.obs,y.pred,thresholds[j])
    }
          kappa.values<-unlist(pcrit[,1])
          sensi<-unlist(pcrit[,2])
          spezi<-unlist(pcrit[,3])
          ccr<-unlist(pcrit[,4])
          diff<-abs(sensi-spezi)
          kappa.max<-max(kappa.values)
          p.kappa.min<-min(thresholds[kappa.values>=max(kappa.values)])
          p.fair.min<-min(thresholds[diff<=min(diff)])
          sensi.k<-min(sensi[kappa.values>=max(kappa.values)])
          spezi.k<-min(spezi[kappa.values>=max(kappa.values)])
          ccr.k<-min(ccr[kappa.values>=max(kappa.values)])
          kappa.f<-min(kappa.values[diff<=min(diff)])
          sensi.f<-min(sensi[diff<=min(diff)])
          spezi.f<-min(spezi[diff<=min(diff)])
          ccr.f<-min(ccr[diff<=min(diff)])

          #Ergebnis, das von der Güte-Funktion zurückgegeben wird
          data.frame(auc,auc.u,auc.o,r2n,p.fair.min,kappa.f,sensi.f,spezi.f,ccr.f,p.kappa.min,kappa.max,sensi.k,spezi.k,ccr.k)
}

#Funktion, damit es piept, wenn's fertig ist
bell <-
  function(N=5)
{
  while(exists(bellfile <- paste("bell", sample(9999, 1), sep = ""),
              where = 1)) {}
  on.exit(remove(bellfile))
  assign(bellfile, rep("\007",N), where = 1, immediate = T)
  commandString <-
    paste("dos(\"type .\\\\.Data\\\\",
          bellfile, "\", out = F, min = T       )",           sep = "")
  eval(parse(text = commandString))
}
#################################################################
guete<-rbind(guete.logistf(obs,pred1),guete.logistf(obs,pred2))
guete.out<-cbind(c(art,art),c(1,2),guete)
exportData(guete.out, Dateiname.guete, type="EXCEL",colNames=T)

#bell()
```

Appendix C. Maps

App. C-1. Rookeries (Rook *Corvus frugilegus* colonies) mapped in 2006.



Appendix D. Habitat photographs

App. D-1. It‘s a curse as well as a blessing: Domestic livestock herds create short swards suitable for Sociable Lapwing breeding, but have been made responsible for egg-trampling, too.

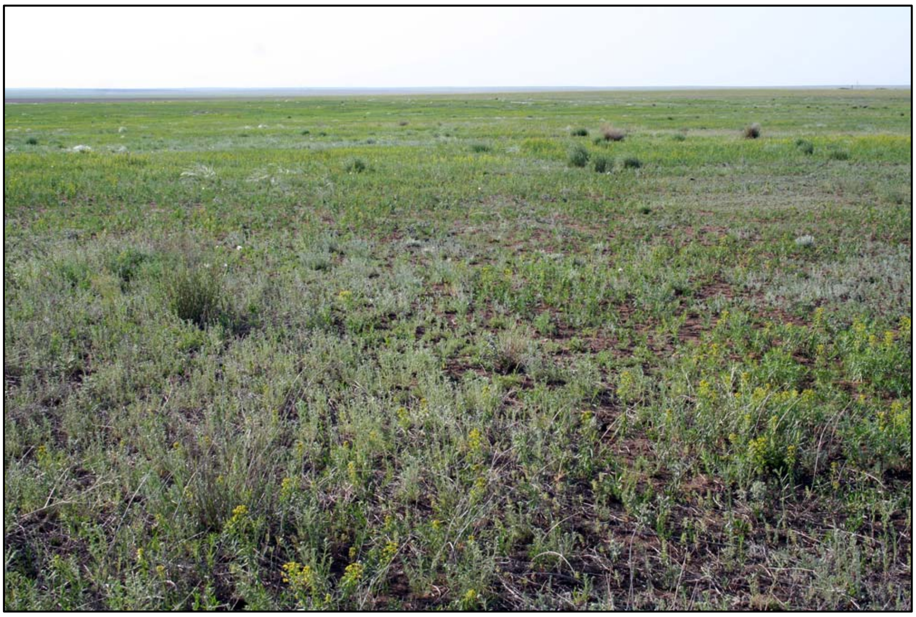
Birlik, Korgalzhynskii raion, Akmolinskaya oblast‘, Kazakhstan. June 2006. Foto: M. Koshkin

App. D-2. Tall, uniform, and virtually ungrazed stands of feather grass (here *Stipa lessingiana)* are avoided by Sociable Lapwing.

Kaskatau, Korgalzhynskii raion, Akmolinskaya oblast‘, Kazakhstan. Mai 2006. Foto: J. Kamp

App. D-3. Sociable Lapwing breeding habitat: uniform, intensively grazed short-grass steppe

Kachiry, Kashyrskii raion, Pavlodarskaya oblast‘, Kazakhstan. Mai 2007. Foto: J. Kamp



App. D-4. Sociable Lapwing breeding habitat: fallow wheat field.

Maishukyr, Korgalzhynskii raion, Akmolinskaya oblast‘, Kazakhstan. Mai 2006. Foto: J. Kamp

App. D-5. Sociable Lapwing breeding habitat: recently burnt long-grass steppe

Maishukyr, Korgalzhynskii raion, Akmolinskaya oblast‘, Kazakhstan. Mai 2006. Foto: J. Kamp

App. D-6. Adult male Sociable Lapwing in full breeding plumage

Aktubek, Nurinskii raion, Karagandinskaya oblast‘, Kazakhstan. Mai 2006. Foto: M. Koshkin

App. D-7. Suitable short-grazed habitat is often available very close to settlements.

Aktubek, Nurinskii raion, Karagandinskaya oblast‘, Kazakhstan. Mai 2006. Foto: M. Koshkin

App. D-8. Sociable Lapwing nest situated in cattle dung pile.

Zhanteke, Korgalzhynksii raion, Akmolinskaya oblast‘, Kazakhstan. Mai 2006. Foto: J. Kamp

App. D-9. Sociable Lapwing nest situated in *Artemisia* sward

Zhaskairat, Kashyrskii raion, Pavlodarskaya oblast‘, Kazakhstan. Mai 2007. Foto: A. Salemgareev

App. D-10. Recently hatched Sociable Lapwing chick at the nest site.

Kulanotpes, Nurinskii raion, Karagandinskaya oblast‘, Kazakhstan. June 2006. Foto: J. Kamp

App. D-11. Adult female Sociable Lapwing, guiding chicks in *Artemisia* vegetation.

Shalkar, Nurinskii raion, Karagandinskaya oblast‘, Kazakhstan. June 2006. Foto: M. Koshkin