Grimm-Seyfarth, A. (2022):
Environmental and training factors affect canine detection probabilities for terrestrial newt surveys
J. Vet. Behav. **57** , 6 - 15

http://dx.doi.org/10.1016/j.jveb.2022.07.013

Environmental and training factors affect canine detection probabilities for terrestrial newt surveys

Annegret Grimm-Seyfarth

PII: S1558-7878(22)00088-0
DOI: https://doi.org/10.1016/j.jveb.2022.07.013
Reference: JVEB 1511

To appear in: *Journal of Veterinary Behavior*

Received date: 25 November 2020
Revised date: 16 June 2022
Accepted date: 30 July 2022

canine amphibian detection probability

Environmental and training factors affect canine detection probabilities for terrestrial newt surveys

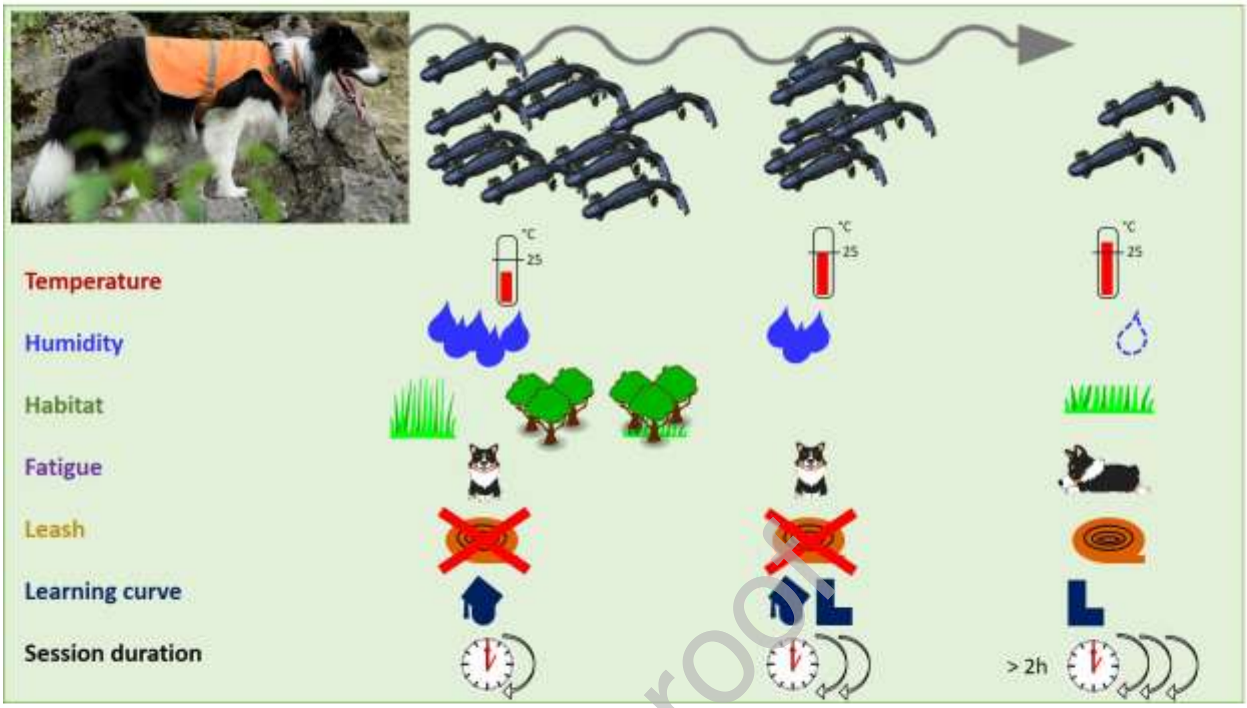
Annegret Grimm-Seyfarth[a,b,*]

[a]UFZ – Helmholtz Centre for Environmental Research, Department of Conservation Biology, Permoserstr. 15, 04318 Leipzig, Germany

[b]Wildlife Detection Dogs e.V., Immenkorv 19, 24582 Bordesholm, Germany

[*]**Corresponding author:** Dr. Annegret Grimm-Seyfarth, Helmholtz-Zentrum für Umweltforschung UFZ: Helmholtz-Zentrum fur Umweltforschung UFZ, Conservation Biology, Leipzig, Sachsen, Germany, E-mail: annegret.grimm@ufz.de

Graphical abstract



Abstract

Amphibians are of great interest to scientific research, but many populations are highly threatened and declining worldwide. Although varieties of aquatic survey methods exist, traditional methods for terrestrial surveys are very time-consuming but often not very effective. A novel method to detect terrestrial amphibians is the use of wildlife detection dogs. While their use for and factors affecting detection rates of mammals and reptiles are well documented, scientific literature on amphibian detection dogs is just emerging. The purpose of this study was to investigate the effects of environmental (habitat, weather) and training factors on detection probabilities for a newt detection dog. An experienced wildlife detection dog was trained and tested on smooth (*Lissotriton vulgaris*) and great crested newts (*Triturus cristatus*). Environmental and training parameters were recorded for 101 test trials and

used as explanatory parameters in a binomial GLM. I found that detection probability strongly depended on temperature, whereby optimal temperatures varied by habitat. Detection probabilities were lowest in short grass, but there was no strong difference among forest habitats. They were higher for males than females and for great crested than smooth newts. For this dog, detection probabilities were also higher if the dog was working off the leash than on it, and when the dog was cooperative than fatigued. Dog performance increased over time with a strong increase at the beginning and a plateau at 92% detection probability. However, detection rates of this specific dog slightly decreased when the dog was working more than two hours. The findings of this study provide a valuable basis for future deployments of this and other amphibian detection dogs. Dogs can certainly work in a variety of different habitats, although directed off-leash searching with enough time in complex habitats and specific training for small species with low detection distances may enhance their performance. The study design might also consider the temperature and humidity at which the dog will be deployed. A regular assessment of the detection dog using blind tests will give an indication on its reliability. Assessments similar to this study may further be used to estimate detection probability for a particular dog under given field conditions. Regular blind tests will show when the detection rates reached a plateau, which may then give an indication on its reliability. Altogether, results suggest that newt detection dogs may provide a highly promising survey method, which can certainly be transferred for detecting other amphibians in terrestrial habitats.

Keywords: *conservation dogs, detection probability, Lissotriton vulgaris, species monitoring, Triturus cristatus, wildlife detection dogs*

Introduction

Due to their fascinating biology and the global threat they face, amphibians are of great interest to scientific research. Many different factors are responsible for the rapid population declines of this animal class, including habitat destruction and fragmentation, pollution, global climate change, the spread of diseases, illegal trade, and the introduction of non-native species that either directly threaten them or represent a new food competition (Hayes et al., 2010; Falaschi et al., 2019). The effect of these factors may be multiplied, as many amphibian species require freshwater to breed and therefore migrate to breeding ponds in spring and back to terrestrial habitats in summer or autumn (Temple and Cox, 2009). In addition to scientific research, amphibians are frequent target groups in management and landscape planning, on the one hand to detect the presence of rare species, and on the other hand to catch and transfer individual animals during construction measures and along roads (Hachtel et al., 2017).

A variety of different methods are available for catching and monitoring amphibians (Heyer et al., 1994; Henle and Veith, 1997; Hachtel et al., 2009). Most of them refer to aquatic surveys during their reproduction period, where calling records and specific water traps can provide very good catch numbers (e.g., Gunzburger, 2007; Drechsler et al. 2010). However, the search on land is much more complicated. Terrestrial traps are usually only effective with an elaborately designed fence and only provide meaningful capture numbers during migrations between water and land habitats (Hill et al., 2005). Outside the migration period, visual encounter or transect surveys are often conducted. They depend on the activity time of the respective species and are very time-consuming but not very effective (English Nature, 2001; Hill et al., 2005; Ali et al. 2018), except for highly mobiles species that are active in open forests and salamanders that can be found when being active after the rain (Heyes et al., 1994). More elusive amphibian species are not sufficiently detectable by this method (Hill et al., 2005). Even in areas with high population densities, only certain species can be detected with most terrestrial monitoring methods, especially in denser habitats such as forests (Vences et al., 2008).

A novel method to detect hidden, burrowing or inactive amphibians is the use of wildlife detection dogs (Powers, 2018; Grimm-Seyfarth and Harms, 2019). Wildlife detection dogs provide a method to monitor species of all kingdoms that could otherwise not or hardly be studied (Dahlgren et al., 2012; Bennett et al., 2019). While their use for birds and mammals dates back to 1890s and 1940s, respectively, and their use for reptile detection has been known since the 1950s (Grimm-Seyfarth and Harms, 2021), the first study on amphibian detection is much more recent (Sokolov et al., 1990). Although amphibian detection dogs have been regularly used to find invasive species such as cane toads (*Chaunus marinus*) in Australia and New Zealand (Peacock, 2007), scientific literature on their performance is rare (Powers, 2018; Grimm-Seyfarth et al., 2021). In contrast to reptile (Cablk and Heaton, 2006; Savidge et al., 2011) and scat detection dogs (Long et al., 2007; MacKay et al., 2008; Reed et al., 2011), I am not aware of any published study that determined the effects of environmental factors on detection rates of amphibian detection dogs.

The purpose of this study was to examine detection probabilities of a newt detection dog as an example of amphibian detection in the terrestrial realm in blind test settings. I aimed at investigating the effects of (i) different habitats and sites, (ii) different weather conditions, (iii) different target species, and (iv) different training conditions on detection rates. Knowledge of these factors will enable the estimation of detection probabilities during fieldwork.

Materials and methods

Study area and amphibian species

The study took place in the Flora-Fauna-Habitat area “Leipziger Auensystem” (Floodplain system of Leipzig, 51°22'01.9"N, 12°15'32.0"E). Training and testing of the newt detection dog were conducted in the “Papitzer Lehmlachen” (51°22'49.7"N 12°14'33.4"E), a part of the nature reserve “Luppeaue”, in 2018, and in the nature reserve “Burgaue” (51°22'05.6"N 12°16'53.3"E) in 2019 and 2020. In those areas, an annual aquatic amphibian monitoring by means of minnow traps took place, which aimed at catching smooth newts (*Lissotriton vulgaris*) and great crested newts (*Triturus cristatus*) during their reproductive period between April and June for population viability analyses. Each newt was measured, weighed and photographed for individual identification and then released at the point of capture (Mazoschek, 2018). I used some of the captured newts for dog training and testing (see 2.2). I made sure that each individual was only used for a short time and kept moist regardless of the environmental conditions.

I have been granted the nature conservation exception from the prohibitions of § 44 para. 1 no. 1, 2 BNatSchG (Federal Nature Conservation Act) and § 4 para. 1 no. 1 BArtSchV (Federal Species Protection Ordinance) by the respective responsible Lower nature conservation authorities. These allowed me (and my colleagues) to catch and handle native amphibian species.

Detection dog training and testing

The dog used in this study was a private-owned pure-bred male Border Collie called “Zammy”, housed and handled by the author of this study. The dog was born in 2016 and trained as a wildlife detection dog from puppy on (Grimm-Seyfarth et al., 2019). He has been, and still is, intensively working as a scat detection dog for Eurasian otter (*Lutra lutra*) and in a pilot study for the Eastern green lizard (*Lacerta viridis*). In a pre-study in 2018, the dog was additionally trained on swab samples from ten male and eleven female smooth newts as well as five male and eight female great crested newts (Grimm-Seyfarth and Harms, 2019). In the same year, I stopped using swabs due to the uncertainty of which odour is actually attached to them. Instead, the dog was familiarized with living newt individuals caught in aquatic surveys (see 2.1). Due to the dependence on live newts, dog training was restricted to April until June each year. During the first trials in the field in 2018, newts were secured inside small plastic cages (7.5 cm x 18 cm x 3 cm, mesh size 0.8 cm). However, the dog has been trained a passive sit-and-stare alert as a response to finding a newt and did not interact with any individual. In order to minimize visual clues, I therefore eliminated the cages in the following years and instead started training on a five-meter towing leash to warrant the safety of the newts. Moreover, as we initially observed that detection distances for such small targets can be as low as 20 cm, a towing leash can optimize search strategy and coverage (Woollett (Smith) et al., 2014). With advances in training, the dog learned to walk slower and focus on small targets by himself. The leash was preferably used in open habitat to prevent the dog from walking too fast and potentially overlooking a newt, but less in dense habitat where he was guided by dog whistle when searching around large woodpiles or shrubs. However, on-leash and off-leash searches have been conducted in all habitat types for comparisons.

In spring 2019, the dog learned to find newts at various search areas and in various habitats. From that year on, most training sessions were accompanied by at least one blind test setting where the dog had to find an individual of unknown species and sex in various search areas of approx. 20 x 50 meters. To ensue blindness, tests were conducted with the help of an assistant who randomly placed the newt in the search area while both the handler and the dog were not present. The assistant made sure that the newt could not dry out or overheat by putting water and / or a shelter of wood or leaves on it and observing that the hiding place would not suddenly be exposed to direct sunlight. This should have also prevented the newts from roaming around a lot.

He/she also always touched other parts of the search area and walked through it several times to exclude that the dog would only detect human scent instead of the newt itself. Furthermore, seven different assistants alternated to place the newt and occasionally, more than one assistant moved through the search area to spread human scent to prevent the dog from trailing a specific person. After placing the individual, the assistant stood still either in the search area or at the border of it. He/she was instructed to not interact with the dog during the search, nor to look directly towards both the dog and the hiding place of the newt to avoid non-verbal cues to the dog. Every search started with the handler placing the dog at the border of the search area and sending him to search on command. The dog was allowed to search freely wherever possible and called to turn when he left the search area.

In one third of all test settings, the assistant randomly placed between one and four additional decoys in the search area. I define those decoys as anything that could hold newt scent, such as the water or glass in which the newt was kept before, or any tool that had touched it. This was because in later field searches, the dog should only find living individuals or occupied burrows and not every shelter a newt had used on its way through the habitat. Thus, the dog was trained to only detect live newts rather than residual newt odour. The handler never knew whether or how many decoys had been placed. The test stopped either when the dog alerted or when the handler decided that the dog has searched the area and did not find the target. After each alert, the handler approached the dog and looked for the newt. If the dog had alerted at a decoy, the assistant told the handler and picked up the hidden newt. Tests and trainings usually alternated within a session. Each session comprised an average of seven tests (between one and 15 tests) and lasted an average of 1.4 hours (between 0.5 and 2.5 hours), with longer sessions comprising more tests or training trials. I always finished a session with the dog finding a newt and alerting at it. I used positive reinforcement without coercion. For each correct alert, the dog gets a playtime with its ball, followed by a food reward to complete his chain of action for hunting. If the dog alerted at anything other than the living newts during training, including the decoys holding newt scent, I ignored that alert and sent him on searching. After each test and training session, I recorded the species and sex of the hidden newt, the number of decoys, whether the dog alerted at the newt (1/0) or a decoy (1/0), whether the dog was leashed, the condition of the dog, and the duration of the whole session (Table 1). A total of 101 blind tests have been conducted, 44 in 2019 and 57 in 2020.

Environmental parameters

Throughout testing and training in the field settings, the dog was confronted with a range of environmental conditions, all of which could influence detection probabilities (MacKay et al., 2008). I recorded the temperature and relative humidity (TFA Dostmann handheld devices), the general weather and wind condition, and the flying insect activity. I further recorded the predominant terrain and its elevation of the search area. Last, I noted the specific site where the newt was hidden and whether it was placed visible or not (Table 1).

Statistical analyses

I defined the naïve detection rate as the proportion of detected individuals. Applied to the detection dog method, this meant the number of true-positive alerts divided by the number of targets hidden. This is also often referred to as sensitivity of a detection dog. I also calculated the dog's precision as the number of true-positive alerts divided by the total number of alerts, i.e. including all false-positive alerts. False-positive alerts would be alerts at anything other than the living individual, including alerts at the decoys holding newt scent.

The 101 blind tests were used to estimate detection probabilities by means of a generalized linear model (GLM) (Nelder and Wedderburn, 1972) with binomial error distribution, whereby "1" implies a correct and "0" no alert at the hidden newt per test. Training (i.e., parameters that can be manipulated during training) and environmental parameters (i.e., parameters that depend on given environmental conditions) were used as explanatory parameters, together with the test number (i.e., the chronological numbering of tests). In order to avoid collinearity, I performed correlation tests among all mentioned parameters beforehand. (Dormann et al., 2013). Since many parameters showed slight correlations among each other, I chose a conservative correlation threshold of $R \geq 0.4$. Following this, I removed the parameters relative humidity, wind, elevation, hidden, and insects (Appendix S1). I also removed the weather since most tests had been conducted on partly cloudy days (77.2%), in comparison to 5.9% sunny, 8.9% cloudy and 7.9% drizzling days. I then built a model with the remaining parameters without interactions and calculated the variance inflation factor using the R package "car" (Fox and Weisberg, 2019). All

GVIF^1/2df (Fox and Monette, 1992) were below 5 and thus did not indicate further serious collinearity problems.

The final global GLM included the training parameters *Test number, on/off Leash, Session duration, Species, Sex, Number of decoys* and *Condition of the dog*, and the environmental parameters *Site, Temperature, Terrain*, and the interaction between the latter. As the temperature could show an optimum curve for the detection probability, I compared the model to one with a quadratic relation for temperature using Akaike's Information Criterion corrected for small samples (AICc). Likewise, I considered that the dog's performance could show a saturation curve over time and compared the model to one with a logarithmic relation to test number. After selecting the global model, I fitted all possible model combinations within the given set of test predictors (Stephens et al., 2006) using the R package "MuMIn" (Barton, 2019) and compared them using AICc (Burnham and Anderson, 2002).

I obtained parameter significances by means of a likelihood ratio test (LRT) of the full model against the model without the parameter in question and overall model significance using an LRT of the full model against a model including the intercept only. I obtained parameter estimates and standard errors from the full model (Cade, 2015) and their relative importance by summing up their AICc weights (ωAICc) across all models. The model did not show any visual sign of autocorrelation. All analyses were performed in R (R Core Team, 2020).

Results

Summary of raw data

Of the 101 blind tests conducted, 45 (16 females, 29 males) refer to the smooth newt and 56 (28 females, 28 males) to the great crested newt. In 2019, great crested newts were dominant (31 *T. cristatus* vs. 13 *L. vulgaris*) and the overall sex ratio was equal (22:22), while in 2020, smooth newts were dominant (25 *T. cristatus* vs. 32 *L. vulgaris*) and the sex ratio was slightly male-biased (35:22). Around two third (67) of the tests have been conducted without decoys, 23 with one decoy, six with two, four with three and one with four decoys. More tests have been conducted on leash (62) than off (39). The dog was mostly cooperative (83), rarely fatigued (18), and never distracted or frustrated. Total test and training sessions lasted from 30 to 150 minutes with an average of 102 and a median of 120 minutes.

Tests have been conducted at temperatures from 9.9°C to 27.7°C with an average of 19.64°C, and relative humidity from 41.8% to 80.8% with an average of 57.61%. Higher temperatures were strongly correlated with lower humidity ($R = -0.59$, $p << 0.001$) and a higher amount of flying insects ($R = 0.69$, $p << 0.001$). Despite the weather was partly cloudy most of the test days, it also correlated with temperature ($R = 0.31$, $p << 0.001$). Higher temperature also increased the occurrence of fatigued condition in the detection dog ($R = 0.39$, $p << 0.001$) (Fig. 2, Appendix S1). Test terrains were almost equally distributed with 21 tests in short grass, 18 in tall grass, 30 in open understory and 32 in dense understory. Most newts were placed in grass / litter (60), followed by on / under logs (37), on tree trunks (3) and under rocks (1).

All newts were found again. They usually stayed in place and only rarely moved up to one meter. The total naïve detection rate (i.e., the percentage of newts detected by dogs in relation to the total number of placed newts) of newts was 77.23%. Separated by year, the naïve detection rates in 2019 and 2020 were 65.91% and 85.96%, respectively. Of the 34 tests where decoys had been placed, the total naïve detection rate of newts was 82.35% and false-positive alerts at decoys amounted to 11.74% of the tests, but only 7.84% (four decoys) of all decoys placed. Separated by year, the naïve detection rates in 2019 and 2020 with decoys present were 50% and 92.31%. False-positive alerts at decoys amounted to 37.5% of the tests and 25% (three decoys) of all decoys placed in 2019, as well as 3.85% of all tests and 2.56% (one decoy) of all decoys placed in 2020. The overall precision was 95.12% throughout all tests, 90.63% in 2019, and 98% in 2020. No false positive alert occurred on other objects than the decoys holding newt scent.

Statistical analyses

I rejected the quadratic relation to temperature as this weakened the model ($AICc_{quadratic.temperature}$ = 122.4, $AICc_{linear}$ = 112.1) but kept the logarithmic relation to test number as this slightly improved it ($AICc_{log.test.number}$ = 110.7). The global model was significantly different from the null model (LRT, p = 0.0007).

The best model explaining the influence on detection probability included the training parameter *Sex* and the environmental parameters *Temperature, Terrain*, and their interaction (Table 2). While no further environmental parameter influenced detection probability, all other training parameters were included individually or in combinations of two or three in the best models within ΔAICc < 2 (Table 2). Thus the most important parameters were, in ascending order: *Terrain, Sex, Temperature, Temperature:Terrain, Condition of the Dog, on/off Leash, Number of Decoys, Session duration, Test number*, and *Species* (Table 3).

With increasing temperature, detection probability decreased in short grass and open understory, but increased in tall grass and dense understory. It was higher for males than females and for the great crested newt than the smooth newt (Fig. 3). Detection probability was also higher if the dog was working off the leash than on it, and when the dog was cooperative than fatigued (Fig. 4). Notably, when adding the parameters *Condition* and *on/off Leash* to the model, detection probability decreased with increasing temperature in dense understory (Fig. 4). Detection probability was slightly decreasing with increasing session duration, but increasing with a higher number of decoys and advancing test number (Fig. 5). The average detection rate for test numbers 90-120 was estimated at 0.92. It did not differ substantially among terrains, but was usually highest for tall and lowest for short grass (Fig. 5).

Discussion

Already naïve detection rates for a newt individual on 0.1 ha indicated that the detection dog performed better in 2020 than in 2019. Statistical analyses confirmed the increasing sensitivity, showing a logarithmic increase in estimated detection rates levelling off at 92%. While the overall naïve detection rate was comparable to that of placed California tiger salamander (*Ambystoma californiense*) hybrids on 0.1 ha (79%) (Powers, 2018), final detection rates were similar to overall detection rates for placed individuals of Desert Tortoise (*Gopherus agassizii*) on 0.5 and 2 ha (91%) (Cablk and Heaton, 2006). However, it is possible that the dog would not reach that plateau when searching for wild newts on larger areas, as also detection rates of the Desert Tortoise detection dogs decreased to 70% when searching for wild tortoise on 6.25 ha subplots (Nussear et al., 2008). Nevertheless, the increased accuracy over time was attributed to the learning curve in Desert Tortoise detection dogs (Cablk and Heaton, 2006). Likewise, Taranto (2019) described a very steep linear learning curve at the beginning, while Gompper et al. (2006) assumed that "final" detection dogs show a constant detection rate. Together, this is pictured in a logarithmic learning curve by the newt detection dog in this study with a steep increase in 2019 and a plateau in 2020, indicating that the dog has finished training and may be operational capable. Therefore, regular validation tests until detection rates level off may be one appropriate means to determine operational capability of wildlife detection dogs. Thus, it seems useful to monitor the performance of any wildlife detection dog during training to be able to detect each dog's specific plateau for the task and target under specific conditions.

Short after the number of true-positive indications increases at the beginning of training a detection dog, the number of false-positive indications also increases, but soon declines again (Taranto, 2019). In line with that, in this study the number of false-positive alerts in 2019 was ten times higher than in 2020. Importantly, the dog only made false-positive alerts at the decoys holding newt scent, i.e., material that has touched the newt. It may have not been obvious for the dog to ignore these decoys at the beginning, as he was alerting at scent traces of the correct species, also known as residual odour. It has been shown that detection dogs can recognise the odour of the targeted amphibian species in concentrations as low as 1:100,000 (Matthew et al., 2021). At the beginning, the dog may have additionally been confused as he was pre-trained on newt swabs in 2018 and it remains unclear which scents exactly stick to the swabs and what the dog learned from it. In 2019, he thus had to learn to ignore anything other than the living individuals, including residual odour. This is particularly important for retreat searches where many parts of the habitat may hold newt scent, but only occupied retreats are of interest. The fact that the detection rate of the dog increased with the number of decoys placed indicates that this training was successful. Furthermore, his very high precision in 2020 of 98% indicates that detection dogs can be a highly reliable detection method for living individuals while ignoring residual odour such as previously occupied hiding places still holding newt scent. Other studies found that amphibian detection dogs can exhibit near-perfect specificity when tested against placed non-target species (98.6%, (Matthew et al., 2021) and 100% (Powers, 2018)).

While it seems logical that the detection rate for the larger species (great crested newt) was higher than for the smaller one (smooth newt), as larger species may produce more scents that are easier to detect from a greater distance, the difference was much stronger for the sex than for the species, with higher detection rates for males. As we almost exclusively used newts in their reproductive period for blind tests, these individuals should have also produced courtship pheromones. In male European newts, these pheromones are particularly strong and produced by a specialized abdominal gland (Malacarne and Giacoma, 1986). They also have a scent marking function and are thus likely to be present during the terrestrial phase as well, even though less expressed (Malacarne and Giacoma, 1986). If male newts indeed produce substantially more scent than females, it is likely that the detection dog could smell males from a much greater distance, and larger detection distances may increase detection probabilities, particularly in dense habitats. In addition, sample size in 2020 was slightly biased towards male smooth newts, which may contribute to the higher detection rates for males.

The newt detection dog was trained under a range of field conditions in natural newt habitats. Detection rates strongly depended on temperature and terrain, usually decreasing with advancing temperature, but increasing in tall grass. However, sample size in tall grass was lowest (N = 18) and all tests but those at 10°C yielded an 100% detection rate. Together with the fact that most tests were conducted at above 15°C, this may indicate that detection rates increase up to ca. 12°C, reach a plateau, and decrease above ca. 25°C, whereby exact temperatures may depend on the habitat and on the individual dog. While previous studies did not detect effects of temperature on detection rates for tortoise detection (Cablk and Heaton, 2006; Nussear et al., 2008), effects on bird detection were comparable to those found in this study (Gutzwiller, 1990), and effects on scat detection were dog-dependent (Reed et al., 2011). Importantly, higher temperatures were correlated to more direct sunlight, which can impede scent production and thus, decrease detection probability (MacKay et al., 2008). Higher temperatures were also correlated to lower relative humidity. In turn, increasing relative humidity also increased detection probability, which was also found for snake (Savidge et al., 2011) and scat detection (Reed et al., 2011), as humidity likely enhances scents (Osterkamp, 2020). Higher temperatures also favoured flying insects, and a higher amount of biting insects could detain the dog from staying focussed (MacKay et al., 2008), but they can also form plumes which can move scent molecules and decrease detection probability (Osterkamp, 2020). Lastly, higher temperatures increased fatigue in the dog, which is due to higher panting rates, which in turn reduce olfactory performance (Osterkamp, 2020).

Apart from temperature and humidity, wind is assumed to have a substantial effect on scent detection (MacKay et al., 2008; Reed et al., 2011; Osterkamp, 2020). Regrettably, in this study, wind strength was strongly correlated with test number and was much stronger during tests in 2019. Results are thus confounded with the dog's learning curve. However, I observed that detection probabilities in 2019 on short grass were higher when the dog was working with the wind. It is likely that the wind pushes the scent of the very small target individuals towards the grass leaves behind it, which then reflects the scent particles back towards the dog. When the dog works with the wind, it would get these scent particles and can work out the target, while when the dog works against the wind, scent particles would be distributed away from the dog making it harder to detect the target (Osterkamp, 2020). The effects of environmental and training factors on detection probabilities were stronger in short grass than in forest, and no substantial differences occurred between dense and open understory. It is possible that scents of targets as small as newts becomes less diffused in forests than in short grass. Moreover, as the microclimate in forests is much more stable than in open habitats, scents may get less dispersed (Osterkamp, 2020). Other studies also found low differences among habitat types, but search duration may increase in complex habitat (Leigh and Dominick, 2015).

Although effects were small, detection probabilities were lower when the dog was fatigued than when he was cooperative, and when he searched on-leash than when he searched off-leash. They were particularly low when the dog was both leashed and fatigued and maybe less willing to persuade the handler to follow him. Generally, it is believed that working off-leash maximises searching ability (MacKay et al., 2008), may enable greater coverage of the area (Woollett (Smith) et al., 2014), and may allow dogs to compensate wind conditions (Reed et al., 2011). For example, in a koala (*Phascolarctos cinereu*s) scat detection dog, false negatives only occurred when the dog was leashed (Cristescu et al., 2015). In this study, the effect was particularly strong on short grass, where scents dispersed stronger and wind was probably more important. This is particularly interesting as we started the training of the dog on-leash, as it is recommended for small targets with low detection distances (Woollett (Smith) et al., 2014). However, a dog that is to be deployed in dense habitats, like the newt detection dog in this study, can hardly be kept on a leash (MacKay et al., 2008, Woollett (Smith) et al., 2014). Therefore, a training prerequisite when searching small targets in dense habitats is that the dog has to learn to search for small targets independently while being directed by the handler from a distance. Results of this study suggest that a dog trained in this way may end up being better able to find the target when searching off-leash. Nevertheless, a final decision will be both project- and dog-dependent (MacKay et al., 2008). Lastly, detection probability decreased with session duration, but the effect was lower than that of the other factors. Nevertheless, it indicates that fieldwork might need longer breaks every two hours to maximise success, but further research is needed based on true field deployments.

The use of only one dog is an important point for the interpretation of the results of this study. Regrettably, most such studies have only one or two dogs available (e.g., Reed et al., 2011; Savidge et al., 2011; Leigh and Dominick, 2015). However, the dog used is an experienced wildlife detection dog that has undergone rigorous training and evaluation for its new target species. With the high test number under a variety of training and environmental conditions, a final and stable detection rate of 0.92, and general findings being supported by other published studies, I believe that these results give a sound indication of the effects of training and environmental factors on detection probabilities for other amphibian detection dogs. Although further studies on this subject with more dogs, amphibian species and conditions are highly recommended, I would believe that they would create more variation among results, but would not change the general findings and significance of this study. Other studies with two dogs found only slight differences among final trained wildlife detection dogs for tortoise and mammal scats (Cablk and Heaton, 2006; Reed et al., 2011), and high detection rates are rather depending on a strong dog-handler-foundation (DeMatteo et al., 2019).

The findings in this study showed that the use of an amphibian detection dog is possible in a range of environments and temperatures and detection rates above 90% can be expected in controlled trials. With the exceptions of hot temperatures (ca. 30°C) in short grass and a fatigued leashed dog, detection probability for this dog was always estimated to be at least 40%. That means that the optimum number of surveys to conduct at each 20 x 50 meters site would vary between three and seven times, depending on the occupancy probability of the site (MacKenzie and Royle, 2005). In comparison, the English Nature states that 60 trapping nights and refuge searches are needed to provide reliable information on the presence and absence of great crested newts (English Nature, 2001). This clearly stresses the suitability of amphibian detection dogs in terrestrial surveys, as other traditional methods have difficulties in detecting those cryptic animals.

Conclusion

Amphibian detection dogs can work in a variety of different habitats. However, special training for detecting small species with low detection distances in complex habitat through directed searches is required. Then, directed searching off-leash may maximise detection probability. The study design might also consider the temperature and humidity at which the dog will be deployed. In short grass, wind conditions affect detection probability most, but avoiding strong wind and searching with the wind will likely circumvent these difficulties. A regular assessment of the detection dog using blind tests will show when the detection rates reached a plateau, which may be one method to give an indication on its reliability. Assessments similar to this study can further be used to estimate detection probability for a particular wildlife detection dog under given field conditions. Together, amphibian detection dogs provide a promising survey method for detecting amphibians in terrestrial habitats.

EthicsStatement

The *Journal of Veterinary Behavior: Clinical Applications and Research* encourages submission of multi-author papers and those with acknowledgments that accurately reflect help received in the preparation of the manuscript or in the research and analysis.

Because of ethical issues raised by recent scientific debate, we wish for the authors to note that we do not accept papers that have “courtesy” authorships, nor those where acknowledgments do not accurately reflect contributions. The following are not acceptable:

- Duplicate (Double) submission
- Submission of the same paper to more than one journal while a decision from another journal on that same paper is still pending
- Repetitive (Redundant) submission
- Reporting the same results or methodologies in somewhat different form
- Improper authorship
- Crediting individuals who did NOT provide a substantive contribution to the research and the analysis presented in the paper
- Lack of credit to individuals who DID provide a substantive contribution
- Lack of conflict of interest disclosure
- Failure to adhere to methodologies consistent with guidelines that may involve the treatment, consent, or privacy of research or testing subjects

To ensure that these requirements are understood and met, we ask that the senior author sign below and upload the signed statement to EES along with the manuscript and any other material, and that they provide us with the e-mail addresses of all other authors and anyone named in the acknowledgments. We thank you for your understanding of and compliance with this necessary detail. Sincerely, Editor-in-Chief

Supporting information

Appendix S1: Correlation plot of all environmental and training parameters. Appendix S2: Test data, original dataset.

Declaration of Competing Interest

The author has declared that no competing interests exist.

Data statement

All test data can be found in Appendix S2.

Acknowledgements

I would like to thank all helpers during test trials, in particular Wiebke Harms and Veronika Koch. I would also like to thank Linda Mazoschek for keeping newts for training while conducting aquatic surveys. Parts of this research have been paid through the Helmholtz International Fellow Award, grant number DP-615, RA-485/19.

References

Aiken, L.S., West, S.G., 1991. Multiple regression: Testing and interpreting interactions. Sage, Newbury Park, USA.

Ali, W., Javid, A., Bhukhari, S.M., Hussain, A., Hussain, S.M. and Rafique, H., 2018. Comparison of different trapping techniques used in herpetofaunal monitoring: A review. Punjab Univ. J. Zool. 33, 57-68. http:dx.doi.org/10.17582/pujz/2018.33.1.57.68

Barton, K., 2019. MuMIn: Multi-Model Inference.

Bennett, E.M., Hauser, C.E., Moore, J.L., 2019. Evaluating conservation dogs in the search for rare species. Conserv. Biol. 34, 314–325. https://doi.org/10.1111/cobi.13431

Burnham, K., Anderson, D., 2002. Model Selection and Multimodel Inference : A Practical Information-Theoretic Approach. Springe US, New York.

Cablk, M.E., Heaton, J.S., 2006. Accuracy and reliability of dogs in surveying for desert tortoise (Gopherus agassizii). Ecol. Appl. 16, 1926–1935. https://doi.org/10.1890/1051-0761(2006)016[1926:AARODI]2.0.CO;2

Cade, B., 2015. Model averaging and muddled multimodal inferences. Ecology 96, 2370–2382. https://doi.org/10.1890/14-1639.1

Cristescu, R.H., Foley, E., Markula, A., Jackson, G., Jones, D., Frère, C., 2015. Accuracy and efficiency of detection dogs: A powerful new tool for koala conservation and management. Sci. Rep. 5, 8349. https://doi.org/10.1038/srep08349

Dahlgren, D.K., Elmore, R.D., Smith, D.A., Hurt, A., Arnett, E.B., Connelly, J.W., 2012. Use of dogs in wildlife research and management, in: Silvy, N.J. (Ed.), The Wildlife Techniques Manual. The John Hopkins University Press, Baltimore, pp. 140–153.

DeMatteo, K.E., Davenport, B., Wilson, L.E., 2019. Back to the basics with conservation detection dogs: fundamentals for success. Wildlife Biol. 2019, wlb.00584. https://doi.org/10.2981/wlb.00584

Dormann, C.F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., Marquéz, J.R.G., Gruber, B., Lafourcade, B., Leitão, P.J., Münkemüller, T., Mcclean, C., Osborne, P.E., Reineking, B., Schröder, B., Skidmore, A.K., Zurell, D., Lautenbach, S., 2013. Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. Ecography (Cop.). 36, 27–46. https://doi.org/10.1111/j.1600-0587.2012.07348.x

Drechsler, A., Bock, D., Ortmann, D., Steinfartz, S., 2010. Ortmann's funnel trap – a highly efficient tool for monitoring amphibian species. Herpetol. Notes 3, 13-21.

English Nature, 2001. Great crested newt mitigation guidelines working today for nature tomorrow. Enshlish Nature, Peterborough, England.

Falaschi, M., Manenti, R., Thuiller, W., Ficetola, G.F., 2019. Continental-scale determinants of population trends in European amphibians and reptiles. Glob. Chang. Biol. 25, 3504–3515. https://doi.org/10.1111/gcb.14739

Fox, J., Monette, G., 1992. Generalized collinearity diagnostics. J. Am. Stat. Assoc. 87, 178–183. https://doi.org/10.1080/01621459.1992.10475190

Fox, J., Weisberg, S., 2019. An {R} Companion to Applied Regression, Third. ed. Sage, Thousand Oaks, USA.

Gompper, M.E., Kays, R.W., Ray, J.C., Lapoint, S.D., Bogan, D, A., Cryan, J.R., 2006. A comparison of noninvasive techniques to survey carnivore communities in Northeastern North America. Wildl. Soc. Bull. 34, 1142–1151.

https://doi.org/10.2193/0091-7648(2006)34[1142:acontt]2.0.co;2

Grimm-Seyfarth, A., Harms, W., 2019. Evaluierung von Artenspürhunden beim Monitoring von Amphibien und Reptilien. Jahresschrift für Feldherpetologie und Ichthyofaunistik Sachsen 20, 56–69.

Grimm-Seyfarth, A., Zarzycka, A., Nitz, T., Heynig, L., Weissheimer, N., Lampa, S., Klenke, R., 2019. Performance of detection dogs and visual searches for scat detection and discrimination amongst related species with identical diets. Nat. Conserv. 37, 81–98. https://doi.org/10.3897/natureconservation.37.48208

Grimm-Seyfarth, A., Harms, W., Berger, A., 2021. Detection dogs in nature conservation: A database on their world-wide deployment with a review on breeds used and their performance compared to other methods. Methods Ecol. Evol. 12, 568-579.

Gunzburger, M.S., 2007. Evaluation of seven aquatic sampling methods for amphibians and other aquatic fauna. Appl. Herpetol. 4, 47-63.

Gutzwiller, K.J., 1990. Minimizing dog-induced biases in game bird research. Wildl. Soc. Bull. 18, 351–356.

Hachtel, M., Göcking, C., Menke, N., Schulte, U., Schwartze, M., Weddeling, K., 2017. Um- und Wiederansiedlung von Amphibien und Reptilien: Beispiele, Probleme, Lösungsansätze. Supplemente der Zeitschrift für Feldherpetologie. Laurenti Verlag, Bielefeld, Germany.

Hachtel, M., Schlüpmann, M., Thiesmeier, B., Weddeling, K., 2009. Methoden der Feldherpetologie. Supplemente der Zeitschrift für Feldherpetologie. Laurenti Verlag, Bielefeld, Germany.

Hayes, T.B., Falso, P., Gallipeau, S., Stice, M., 2010. The cause of global amphibian declines: A developmental endocrinologist’s perspective. J. Exp. Biol. 213, 921–

933. https://doi.org/10.1242/jeb.040865

Henle, K., Veith, M., 1997. Naturschutzrelevante Methoden der Feldherpetologie, Mertensiella. DGHT, Rheinbach, Germany.

Heyer, W.R., Donnelly, M.A., McDiarmid, R.W., Hayek, L.-A.C., Foster, M.S., 1994. Measuring and monitoring biological diversity. Standard methods for amphibians. Smithonian Institution, Washington, DC.

Hill, D., Fasham, M., Tucker, G., Shewry, M., Shaw, P., 2005. Handbook of biodiversity methods. Survey, evaluation and monitoring. Cambridge University Press, Cambridge, UK.

Isyumov, N., Davenport, A.G., 1975. The ground level wind environment in built-up areas. Proc. 4th Int. Conf. Wind Effcts on Buildings and Structures, London, 420-422.

Leigh, K.A., Dominick, M., 2015. An assessment of the effects of habitat structure on the scat finding performance of a wildlife detection dog. Methods Ecol. Evol. 6, 745–752. https://doi.org/10.1111/2041-210X.12374

Long, R.A., Donovan, T.M., MacKay, P., Zielinski, W.J., Buzas, J.S., 2007. Effectiveness of scat detection dogs for detecting forest carnivores. J. Wildl. Manage. 71, 2007–2017. https://doi.org/10.2193/2006-230

MacKay, Paula, Smith, D.A., Long, Robert A., Parker, M., 2008. Scat detection dogs, in: Long, R.A., MacKay, P., Zielinski, W.J., Ray, J.C. (Eds.), Noninvasive Survey Methods for Carnivores. Island Press, Washington-Covelo-London, pp. 183–222.

MacKenzie, D.I., Royle, A. 2005. Designing occupancy studies: general advice and allocating survey effort. J. Appl. Ecol. 42, 1105–1114. https://doi.org/10.1111/j.1365-2664.2005.01098.x

Malacarne, G., Giacoma, C., 1986. Chemical signals in European newt courtship. Ital. J. Zool. 53, 79–83. https://doi.org/10.1080/11250008609355487

Matthew, E.E., Verster, R., Weldon, C., 2021. A case study in canine detection of giant bullfrog scent. J. Vertebr. Biol. 69, 20043.1-11. https://doi.org/10.25225/jvb.20043

Mazoschek, L., 2018. Auswirkungen gewässerspezifischer Faktoren auf Molchpopulationen in der Luppe-Aue. University of Leipzig.

Nelder, J.A., Wedderburn, R.W.M., 1972. Generalized Linear Models. J. R. Stat. Soc. 135, 370–384. https://doi.org/10.2307/2344614

Nussear, K.E., Esque, T.C., Heaton, J.S., Cablk, M.E., Drake, K.K., Valentin, C., Yee, J.L., Medica, P.A., 2008. Are wildlife detector dogs or people better at finding desert tortoises (*Gopherus agassizii*). Herpetol. Conserv. Biol. 3, 103–115.

Osterkamp, T., 2020. Detector dogs and scent movement. How weather, terrain, and vegetation influence search strategies. CRC Press, Boca Raton, USA.

Peacock, T., 2007. Community on-ground cane toad control in the Kimberley. Invasive Animals Cooperative Research Centre, University of Canberra, Canberra, Australia.

Powers, R.M., 2018. Detection dog as abassadors and field assistants to protect imperiled reptiles and amphibians, in: Richards, N.L. (Ed.), Using detection dogs to monitor aquatic ecosystem health and protect aquati resources. Palgrave Macmillan, Cham, Switzerland, pp. 25-67.

R Core Team, 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Reed, S.E., Bidlack, A.L., Hurt, A., Getz, W.M., 2011. Detection distance and environmental factors in conservation detection dog surveys. J. Wildl. Manage.

75, 243–251. https://doi.org/10.1002/jwmg.8

Savidge, J.A., Stanford, J.W., Reed, R.N., Haddock, G.R., Adams, A.A.Y., 2011. Canine detection of free-ranging brown treesnakes on Guam. N. Z. J. Ecol. 35, 174–181.

Sokolov, V.E., Sulimov, K.T., Krutova, V.I., 1990. Кинологическая идентификация индивидальных запахов в следах жизнедеятельности четырех видов позвоночных [The kinological identification of individual scents in the traces of the vital activities of four vertebrate species]. lzvestija Akad. Nauk SSSR 4, 556–564.

Stephens, P.A., Buskirk, S.W., Martínez del Rio, C., 2006. Inference in ecology and evolution. Trends Ecol. Evol. 22, 192–197. https://doi.org/10.1016/j.tree.2006.12.003

Taranto, V.C., 2019. Improving detection and identification methods for volatile oranic explosives. PhD Thesis, University of Technology Sydney.

Temple, H.J., Cox, N.A. 2009. European red list of amphibians. Office for Official Publications of the European Communities, Luxembourg.

Vences, M., Chiari, Y., Teschke, M., Randrianiaina, R.-D., Raharivololoniaina, L., Bora, P., Vieites, D.R., Glaw, F., 2008. Which frogs are out there? A preliminary evaluation of survey techniques and identification reliability of Malagasy amphibians. Mus. Reg. di Sci. Nat. Monogr. 45, 233–251.

Woollett (Smith), D.A., Hurt, A., Richards, N.L., 2014. The current and future roles of free-ranging detection dogs in conservation efforts, in: Gompper, M.E. (Ed.), Free-ranging dogs & wildlife conservation. Oxford University Press, Oxford, UK, pp. 239–264.

Table 1. Recorded training and environmental parameters, their class and description.

Group	Parameter	Class	Definition
Training	Condition of the dog	categorical	cooperative: dog obviously showed targeted searching with an upright posture and without panting during the search; fatigued: dog showed targeted searching but without an upright posture and / or with panting during the search as well as a slower movement; distracted: dog did not clearly show targeted searching and also showed other behaviour (this could include listening to noises, observing people / dogs or obviously sniffing other traces such as those of other dogs or game); frustrated: dog would stop searching at all
	Number of Decoys	integer between 0 and 4	anything that could hold newt scent, such as the water or glass in which the newt was kept before, or any tool that had touched it
	Duration	numeric between 0.5 and 2.5 hrs	the duration of the whole session with each session comprising between 1 and 15 single tests
	Leash	categorical	on: dog was on a 5-m towing leash; off: search without leash
	Sex	categorical	male or female individual hidden
	Species	categorical	*L. vulgaris* or *T. cristatus* individual hidden
	Test number	integer between 0 and 101	the consecutive numbering of the blind test conducted; up to 15 tests could happen during a session; further training happened between tests

Environmental	Site	categorical	substrate the individual was placed on; grass / litter, dead wood, rock, or trunk
	Temperature	numeric	temperature in °C
	Humidity	numeric	relative humidity in %
	Weather	categorical	sunny: no clouds visible; partly cloudy: clouds visible but cover less than ¾ of the sky; cloudy: clouds cover at least ¾ of the sky; humid: sky is covered with clouds and it is foggy or soil is humid; drizzle: drizzling rain; raining: more than drizzling rain
	Wind	categorical	wind condition checked by scattering some baby powder and observing trees following Isyumov and Davenport, 1975; none, slight, intermittent, moderate, or severe, referring to Beaufort scale 0-1, 2, 3, 4, 5 or more, respectively
	Insects	categorical	activity of flying insects; none, few, or many, referring to no, single, or clouds of individuals, respectively
	Terrain	categorical	predominant terrain of the search area; short grass: meadow with grass less than ca. 10 cm height; tall grass: meadow with grass of ca. 10 cm or higher; open understory: forest with up to ¼ of the search area covered with shrubs; dense understory: forest with more than ¼ of the search area covered with shrub
	Hidden	categorical	no: individual was placed visible; yes: individual was placed invisible
	Elevation	categorical	level: no elevation; slight: elevation throughout the search area; mix: only some parts of the search area with elevation

Table 2. Model comparisons for all model combinations within ΔAICc < 2. The next model showed a ΔAICc of 2.08. df – degrees of freedom; logLik – log Likelihood; ωAICc – AICc weights; # Decoys – number of decoys.

Training predictor	**Environmental predictor**	**df**	**logLik**	**AICc**	**ΔAICc**	**ωAICc**
Sex	Temperature * Terrain	9	-37.34	94.65	0.00	0.11
Sex + Duration + log(Test number)	Temperature * Terrain	11	-34.86	94.69	0.04	0.11
Sex + Leash	Temperature * Terrain	10	-36.41	95.27	0.62	0.08
Sex + # Decoys	Temperature * Terrain	10	-36.43	95.31	0.66	0.08
Sex + Leash + Condition	Temperature * Terrain	11	-35.22	95.41	0.76	0.08
Sex + Duration	Temperature * Terrain	10	-36.62	95.68	1.03	0.07
Sex + # Decoys + Condition	Temperature * Terrain	11	-35.39	95.75	1.09	0.06
Sex + log(Test number)	Temperature * Terrain	10	-36.69	95.82	1.17	0.06
Sex + Condition	Temperature * Terrain	10	-36.74	95.93	1.28	0.06
Sex + # Decoys + Duration	Temperature * Terrain	11	-35.51	95.99	1.34	0.06
Sex + Species	Temperature * Terrain	10	-36.85	96.15	1.50	0.05
Sex + # Decoys + Duration + log(Test number)	Temperature * Terrain	12	-34.37	96.29	1.64	0.05

Sex + Duration + log(Test number) + Species	Temperature * Terrain	12	-34.46	96.47	1.82	0.04
Sex + Condition + Duration + log(Test number)	Temperature * Terrain	12	-34.47	96.48	1.83	0.04

Table 3. Parameter estimates and their standard errors (SE) obtained from the full model, importance values based on whole-model comparisons, and p-values obtained from LRTs of the full model against the model without the parameter in question. Estimates and standard errors of two-factorial predictors refer to the factor in brackets, those of multi-factorial predictors are shown as range. # Decoys – number of decoys; ":" symbolises interaction.

Group	**Predictor**	**Estimate**	**SE**	**Importance**	**p-value**
	(Intercept)	-4.19	7.60	1	0.58
Training	Condition (fatigued)	-1.27	1.08	0.47	0.23
	# Decoys	0.51	0.55	0.42	0.32
	Duration	-1.24	1.03	0.44	0.17
	Leash (on)	-0.47	1.08	0.43	0.66
	Sex (male)	2.04	0.84	0.84	0.008
	Species (*T.*	0.55	0.77	0.27	0.48

	cristatus)				
	log(Test number)	1.01	0.84	0.41	0.17
Environmental	Site	[-20.83; 0.00]	[0.90; 1.79]	0.03	0.99
	Temperature	0.16	0.28	0.80	*
	Terrain	[-37.37; 2.56]	[6.59; 9.09]	0.89	*
	Temperature : Terrain	[-0.12; 3.67]	[0.33; 0.40]	0.70	0.04

*p-value not indicated because it is conditional on another predictor and thus does not have a meaningful interpretation (Aiken and West, 1991).

Figure 1: Newt detection dog Zammy at the start of a blind test (left; Photo: Daniel Peters) and alerting at a great crested newt (right; Photo: Wiebke Harms).

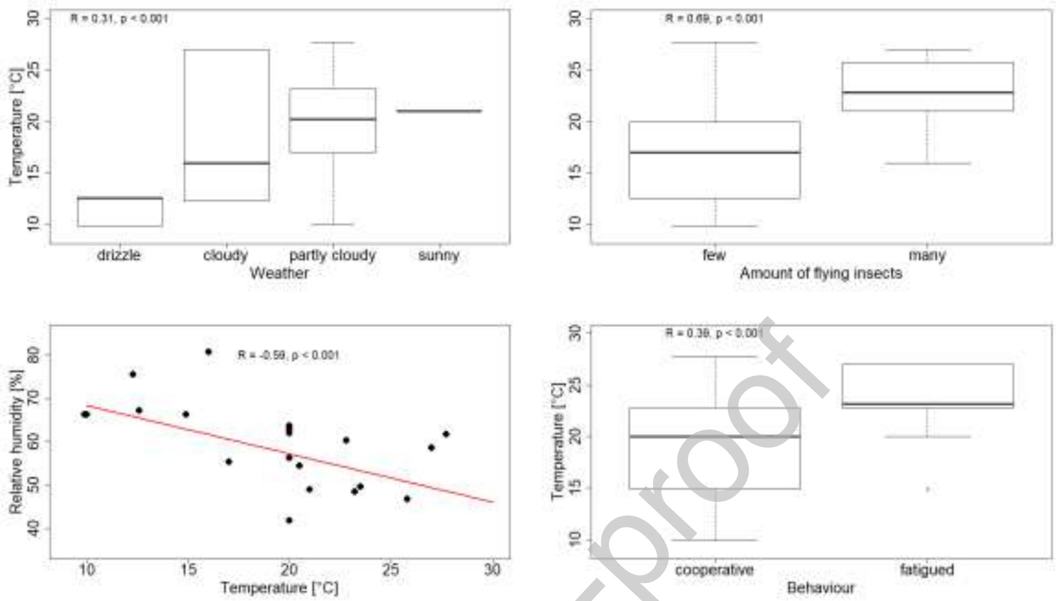


Figure 2: Parameters correlated with temperature. Correlation coefficients (R) and p-values are given in each plot. The red line symbolises the linear relation by means of a linear model.

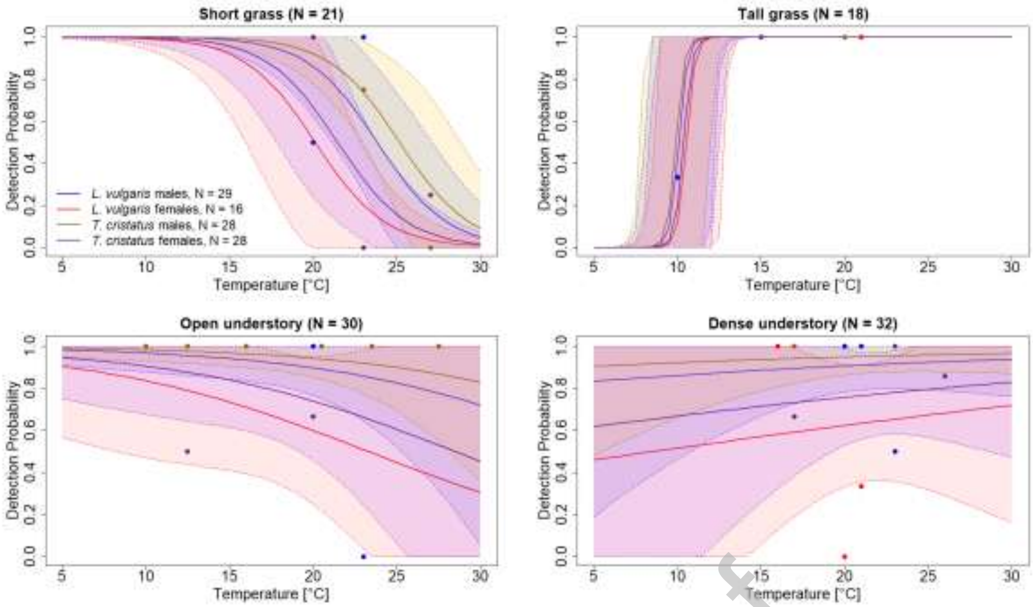


Figure 3: Estimated detection probability in relation to the terrain and temperature, separated (colour-coded) by species and sex. Solid lines represent estimates, dashed lines and light shaded areas their 95% confidence intervals. Dots represent average observed values under a given condition. Sample sizes of observed values per terrain are given above plots, those per species and sex in the legend.

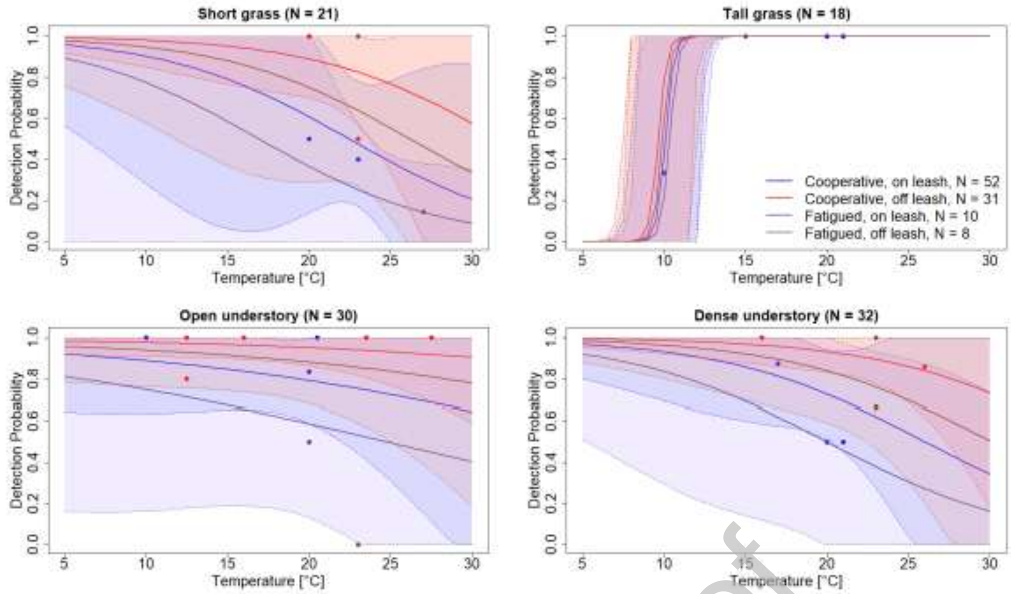


Figure 4: Estimated detection probability in relation to the terrain and temperature, separated (colour-coded) by dog condition and on/off leash. Solid lines represent estimates, dashed lines and light shaded areas their 95% confidence intervals. Dots represent average observed values under a given condition. Sample sizes of observed values per terrain are given above plots, those per condition and on/off-leash in the legend.

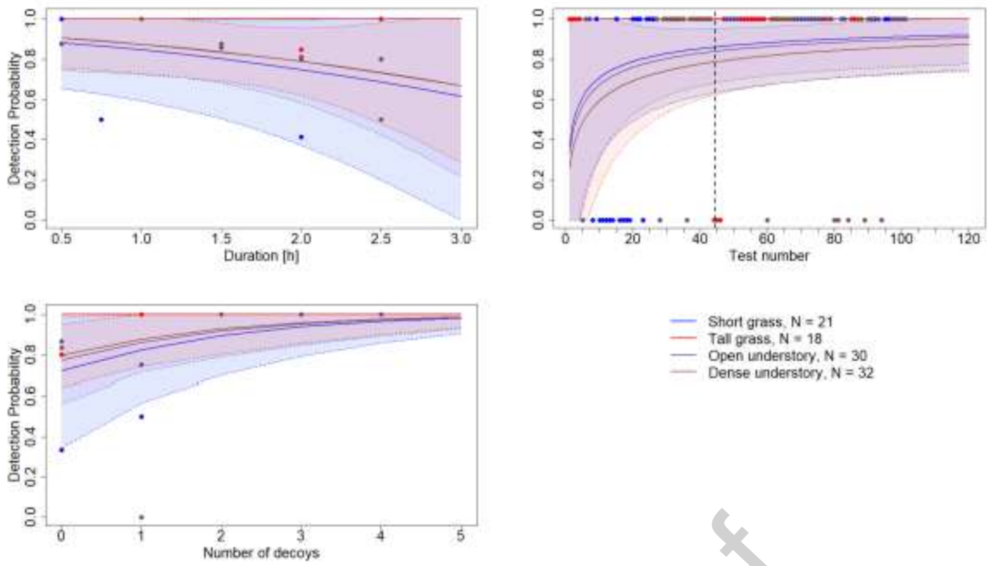


Figure 5: Estimated detection probability in relation to the session duration (upper left), test number (upper right), and number of decoys (lower left), separated (colour-coded) by terrain, for an average temperature of 20°C. Solid lines represent estimates, dashed lines and light shaded areas their 95% confidence intervals. Dots represent average observed values under a given condition but ignoring the temperature. Sample sizes of observed values per terrain are given in the legend. The vertical black dashed line in the upper right plot symbolises the different years, with 2019 left and 2020 right of the line.