

This item is the archived peer-reviewed author-version of:

Motion analysis of non-model organisms using a hierarchical model: Influence of setup enclosure dimensions on gait parameters of Swinhoe's striped squirrels as a test case

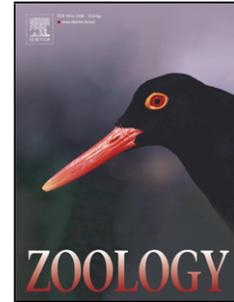
Reference:

Mielke Falk, Schunke Vivian, Woelfer Jan, Nyakatura John A..- Motion analysis of non-model organisms using a hierarchical model: Influence of setup enclosure dimensions on gait parameters of Swinhoe's striped squirrels as a test case
Zoology : analysis of complex systems - ISSN 0944-2006 - 129(2018), p. 35-44
Full text (Publisher's DOI): <https://doi.org/10.1016/J.ZOOL.2018.05.009>
To cite this reference: <https://hdl.handle.net/10067/1548050151162165141>

Accepted Manuscript

Title: Motion analysis of non-model organisms using a hierarchical model: Influence of setup enclosure dimensions on gait parameters of Swinhoe's striped squirrels as a test case

Authors: Falk Mielke, Vivian Schunke, Jan Wölfer, John A. Nyakatura



PII: S0944-2006(17)30252-0
DOI: <https://doi.org/10.1016/j.zool.2018.05.009>
Reference: ZOOL 25648

To appear in:

Received date: 23-11-2017
Revised date: 23-5-2018
Accepted date: 31-5-2018

Please cite this article as: Mielke F, Schunke V, Wölfer J, Nyakatura JA, Motion analysis of non-model organisms using a hierarchical model: Influence of setup enclosure dimensions on gait parameters of Swinhoe's striped squirrels as a test case, *Zoology* (2018), <https://doi.org/10.1016/j.zool.2018.05.009>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Motion analysis of non-model organisms using a hierarchical model:
Influence of setup enclosure dimensions on gait parameters of
Swinhoe's striped squirrels as a test case

Falk Mielke^{1,2#}, Vivian Schunke^{1#}, Jan Wölfer¹ & John A. Nyakatura^{1*}

¹: AG Morphologie und Formengeschichte, Institut für Biologie und "Bild Wissen Gestaltung. Ein interdisziplinäres Labor", Humboldt Universität zu Berlin, Philippstr. 12/13, DE-10115 Berlin, Germany.

²: Applied Veterinary Sciences and Functional Morphology, Universiteit Antwerpen, Universiteitsplein 1, BE-2610 Wilrijk, Belgium.

#: Both authors contributed equally

*: Corresponding author: john.nyakatura@hu-berlin.de

Running title: Hierarchical generalized linear model for motion analysis

HIGHLIGHTS:

- We present an analytical framework to account for the inherent hierarchy of kinematic data.
- The framework uses a hierarchical Bayesian generalized linear model that considers subject specific differences and population effects.
- The hierarchical model can cope with limited sample size and bias introduced by deviating behavior of individuals.
- The framework is exemplified using the test case of an *in-vivo* motion analysis of a non-model organism.
- Our findings justify the common practice of researchers to intuitively select an enclosure with dimensions assumed as “non-constraining”.

ACCEPTED MANUSCRIPT

Abstract

In *in-vivo* motion analyses, data from a limited number of subjects and trials is used as proxy for locomotion properties of entire populations, yet the inherent hierarchy of the individual and population level is usually not accounted for. Despite the increasing availability of hierarchical model frameworks for statistical analyses, they have not been applied extensively to comparative motion analysis. As a case study for the use of hierarchical models, we analyzed locomotor parameters of four Swinhoe's striped squirrels. The small-bodied arboreal mammals exhibit brief bouts of rapid asymmetric gaits. Spatio-temporal parameters on runways with experimentally varied dimensions of the setup enclosure were compared to test for its potentially confounding effects. We applied principal component analysis to evaluate changes to the overall locomotor pattern. A common, non-hierarchical, pooled statistical analysis of the data revealed significant differences in some of the parameters depending on enclosure dimensions. In contrast, we used a hierarchical Bayesian generalized linear model (GLM) that considers subject specific differences and population effects to compare the effect of enclosure dimensions on the measured parameters and the principal components. None of the population effects were confirmed by the hierarchical GLM. The confounding effect of a single subject that deviates in its locomotor behavior is potentially bigger than the influence of the experimental variation in enclosure dimensions. Our findings justify the common practice of researchers to intuitively select an enclosure with dimensions assumed as "non-constraining". Hierarchical models can easily be designed to cope with limited sample size and bias introduced by deviating behavior of individuals. When limited data is available—a typical restriction of *in-vivo* motion analyses of non-model organisms—density distributions of the Bayesian GLM used here remain reliable and the hierarchical structure of the model optimally exploits all available information. We provide code to be adjusted to other research questions.

Keywords:

motion analysis, *Tamiops*, generalized linear model, Bayesian statistics, PyMC3

1. Introduction

In-vivo motion analysis is a widely used method in vertebrate functional morphology and comparative biomechanics to assess basic spatio-temporal parameters and/or specific kinematic or dynamic data of locomotion. Although hypotheses on the functioning of the postcranial musculoskeletal system were traditionally based on an anatomical analysis only (e.g., reviewed in Cosans and Frampton, 2015; Amson and Nyakatura, 2017), motion analyses help to test such hypotheses and were already pioneered in the late 19th century (e.g., Muybridge, 1878; Marey, 1886). Since that time, the number of analyzed species and tools for data analysis is steadily increasing. In functional morphology, however, motion analyses of non-model organisms are most often limited to a small number of subjects. Small sample sizes challenge inference about the population (e.g., the average effects for a given species). Consequently, statistics on data collected from just a few subjects will usually include the repeated exposition of a limited number of subjects with inherently different locomotor capabilities to the same experimental test. In this situation, outlier subjects might bias the results towards a difference in locomotor parameter values that appears significant on the level of the population. This effect can be intensified when sample size differs between subjects, a problem often confronted with when relying on the animals' willingness to cooperate. Similarly, if all subjects are treated separately (i.e., comparing individual averages), limited sampling from the population can be prohibitive for statistical power and models are prone to overfitting. To counteract this, apparently aberrant data (either particular trials or entire data from deviating subjects) are oftentimes excluded from analyses, which might obscure the variability of a species' locomotor characteristics (cf. Andrada et al., 2015). All these issues call for a statistical approach that can integrate all acquired data without losing its conservative character when facing small and biased sample sizes.

Given that a limited number of subjects are repeatedly exposed to the same experimental test and that these subjects are used as proxies for the general locomotor properties of the entire population of a species, the data retrieved from *in-vivo* motion analysis is inherently hierarchical (i.e., data from individuals contribute to population data, which is on a higher hierarchical level). Bayesian hierarchical models (Good, 1980; Lee, 2011) have been suggested to accurately account for the subject animals' individual characteristics as a sample from a probability distribution on population level (Rouder and Lu, 2005; Gelman et al., 2013). Other, non-Bayesian yet hierarchical models (e.g., Generalized Linear Mixed Models, GLMMs) may prove equally valid for the problem of accounting for the inherent hierarchy of motion analysis datasets addressed here (but note that GLMMs do not model distributions for "fixed" effects). For an explanation of the hierarchical nature of usual datasets in motion analyses, please consider the following example: The measurements of contact time of a subject over a series of

trials might be normally distributed. Measuring the same parameter over several subjects can also be described as a higher order distribution from which the subjects' measurements were 'drawn'. Changing any experimental condition potentially introduces individual and also population effects (i.e., effects on both hierarchical levels). Besides the appropriate coverage of hierarchical problems, Bayesian models are flexible and can provide constructions that are customized for a variety of data structures. Hierarchical Bayesian approaches have been successfully established in many fields, for example medicine (Dempster et al., 1983), biodiversity (Iknayan et al., 2014) and ecology (Wikle, 2003). Despite the suggested advantages and the availability of hierarchical models, they have rarely been explicitly applied in the context of *in-vivo* motion analysis (but see related use of linear mixed models, e.g., Duhamel et al., 2004; Vanden Hole et al., 2017).

The majority of laboratory-based *in-vivo* investigations of locomotor behavior aims to study unrestrained locomotor behavior that is ideally identical to the locomotion of the analysed animals' free-ranging kin. Oftentimes, parameters deemed ecologically relevant are systematically varied (e.g., Rollinson and Martin, 1981; Gillis, 1998; Jayne and Ellis, 1998; Bundle and Dial, 2003; Daley and Biewener, 2003; Ashley-Ross and Bechtel, 2004; Franz et al., 2005; Autumn et al., 2006; Lammers et al., 2006; Lemelin and Cartmill, 2010; Schmidt and Fischer, 2010; Shapiro and Young, 2010; Birn-Jeffery and Daley, 2012; Foster and Higham, 2012; Herrel et al., 2013; Birn-Jeffery and Higham, 2014; Chadwell et al., 2015; Youlatos et al., 2015). Transparent enclosures are often used to prevent the subjects from falling, jumping, or flying off the runway (e.g., Schmidt and Fischer, 2011; Nyakatura et al., 2012; Hesse et al., 2015). In other cases, enclosures are a necessary part of the experiment (e.g., respirometry chambers; cf. e.g., Tickle et al., 2012). Nevertheless, little attention has been paid to the potentially confounding effect of the dimensions of the experimental setup enclosure on locomotor behavior. Omitting the statistical assessment of this effect can introduce ambiguity in the assessment of locomotor adaptations to factors the researcher is actually interested in. Dimensions of the setup enclosure can have dramatic effects on posture and locomotion. This is exemplified by the motion analysis of domestic ferrets within a constrained transparent enclosure, which was used to simulate subterranean locomotion (Horner and Biknevicius, 2010; Horner et al., 2016). With this notable exception, we are not aware of systematic tests for the potentially confounding effect of the dimensions of the runway or the runway enclosure in mammal *in-vivo* motion analysis. Oftentimes, the exact dimensions of these experimental setup properties are not even reported. From the experience of one of the authors (J.A.N.) and from our survey of the published literature, enclosure dimensions appear to be decided rather intuitively and are often determined in part due to the conditions of the laboratory room and/or the technical equipment used.

We herein present a study design typical for locomotor investigations, for which the choice between non-hierarchical and hierarchical models made a big difference for the inference on the level of the population. Using a hierarchical Bayesian generalized linear model (GLM; Albert, 1988; Lindley and

Smith, 1972), we analyzed the influence of experimentally varied setup enclosure dimensions (length and height) on locomotor patterns of a non-model organism as a meaningful test case. We compared basic spatio-temporal parameters of the Swinhoe's striped squirrel (*Tamiops swinhoei*, Sciuromorpha, Mammalia)—a small-bodied, agile, arboreal mammal that uses brief bouts of rapid asymmetric gaits (bounds and half-bounds, see below; Fig. 1). Our goal was to demonstrate ramifications of taking the hierarchical nature of data derived from *in-vivo* motion analyses into account by comparing the results of a hierarchical model and a widely used Mann Whitney U test. We studied spatio-temporal parameters of trials representing steady-state locomotion within three enclosures of different dimensions that could well be chosen intuitively by researchers interested in the locomotor characteristics of this species or other non-cursorial, similar sized mammals: The basic enclosure had the dimensions 200 x 20 x 20 cm (length x height x width). A long enclosure (300 x 20 x 20 cm) and a high enclosure (200 x 40 x 20 cm) were assembled by adding modules of similar construction to the basic enclosure. In a longer runway enclosure, more consecutive locomotor cycles are possible and theoretically higher speeds can be achieved. Therefore, we anticipated higher speeds in trials in the long enclosure when compared to the basic enclosure, with longer lasting aerial phases that cover more distance and shorter ground contacts. Similarly, a higher enclosure potentially allows higher bounds. Hence, we hypothesized increased height and duration of the aerial phases compared to the basic enclosure. In addition to the analysis of a number of spatio-temporal parameters, we also used principal component analysis (PCA) to evaluate how the overall locomotor pattern is affected by the experimental manipulation of the enclosure dimensions.

We used a hierarchical Bayesian GLM with Markov-Chain-Monte-Carlo (MCMC) sampling implemented in the freely available Python software toolkit “PyMC3” (Salvatier et al., 2016; all code and data are made available, see Mielke et al., 2018) for data analysis. As we show, this model can accurately incorporate the experimental design (sampling four individuals as representatives of a population) and its influence on the measured parameters and the principal components. Due to its hierarchical characteristic, we demonstrate this approach to be appropriate for typical experimental paradigms that involve intra-group variability and how it can improve inference if sample size is limited or unbalanced. While our study puts emphasis on the benefits that arise from adequately representing the hierarchical structure of the data, the Bayesian character of our proposed approach allows for the display of parameter distributions, even on population level, benefitting from the elegant method of MCMC sampling. Finally, the particular model structure and framework we demonstrate herein is robust to outliers and flexible (e.g., in principle allowing for the direct modelling of non-linear relationships).

2. Materials and Methods

This study involved live animals. All procedures strictly adhered to the pertaining regulations of animal welfare in Berlin, Germany, and were approved by the responsible authorities (LAGeSo reg. no.: 0098/16).

Four young, adult, male Swinhoe's striped squirrels were obtained from a local breeder (Table 1). The animals were kept in a large cage (300 x 200 x 120 cm) at the animal keeping facility of the Institute of Biology at the Humboldt University, Berlin, Germany, and were taken to the experimental setup on recording days. The cage was equipped with several branches and hiding places for each subject. Food and water was available to the animals *ad libitum*. Temperature, humidity, and light conditions were regulated automatically. None of the animals displayed any peculiarities during locomotion (such as limping or hyperventilation).

During each week of a two month period, from October to December 2016, three consecutive recording days were alternated with four subsequent days of rest. Prior tests indicated that an animal's willingness to run within the enclosure was raised when two individuals were placed into the experimental setup at the same time. By this, the animals motivated each other to bouts of locomotion rendering any external motivation (e.g., presenting food or clapping of hands to induce a bout) unnecessary. Recording was done opportunistically whenever animals crossed the field of view (see below). When showing signs of fatigue (squatting and panting), animals were substituted and taken back to their cage for rest. Daily recording sessions never lasted longer than four hours, with the total duration of motion analysis for individual subjects having been much shorter. We used only one setup with experimentally varied enclosure dimensions during each recording day to avoid effects of fatigue and changed the order of tested enclosure dimensions each week.

Highspeed videos were recorded from a lateral perspective using a CamPhF 2000 camera (Photonfocus AG, Lachen, Switzerland) with a framerate of 400 frames per second (fps) and a shutter of 600 Hz covering a field of view of 104 x 35 cm at the central area of the acrylic glass enclosure. The camera was controlled using the software Contemplas Templo (Contemplas GmbH, Kempten, Germany), which also allowed to review videos prior to analysis and management of all videos in a database. The subjects exclusively used asymmetric gaits (bounds and half bounds), which are characterized by both limbs of a girdle moving approximately in phase (Fig. 1). In bounds, forelimbs and hind limbs each touch down simultaneously, whereas in half-bounds forelimbs touch down after one another and hind limbs touch down simultaneously (cf. Hildebrand, 1977, 1989). Each bounding cycle consists of a contact phase (at least one limb in contact with the ground) and an aerial phase (all limbs in the air after fore- and hind limb lift-off). All bouts with obvious acceleration, braking, or turning of the subject were discarded immediately. Segments of individual locomotor bouts (trials) were then analyzed using the software Vicon Motus 3D (Contemplas GmbH, Kempten, Germany). After calibration of the field of view, a representative ('focal') bounding cycle from the middle of the running sequence was selected and the position of the tip of the snout was digitized to assess locomotor speed.

By selecting a bounding cycle from the middle of the bout, we intended to retrieve parameter values close to the maximal speed that the animal used in each setup enclosure. As is typical for non-cursorial animals (Kramer and McLoughlin, 2001; Biknevičius et al., 2013; Andrada et al., 2015), the squirrels displayed intermittent locomotor activity with interspersed pauses and rapid accelerations and decelerations. In order to obtain parameters representative for steady-state locomotion, the aerial phase previous to the contact phase of the focal cycle was compared to the focal aerial phase. The average speed during the first 25% frames of the two aerial phases was compared. In case speed during the faster aerial phase exceeded speed of the slower one by more than 25%, the running sequence was considered non-steady and discarded. It was ensured that each of the two consecutive aerial phases was at least 10 frames long to eliminate consecutive short hops from the analysis, which were sometimes used during exploratory behavior. For all trials meeting these criteria, only the focal locomotor cycle was used to retrieve the characteristic parameters that we used to quantify locomotion. For each of the four subject animals, we aimed to include 20 trials in the final analyses. The sample size was intentionally balanced (see below) and limited by one single subject/condition combination for which only 18 measurements could be obtained. In total, 238 trials were analyzed (Table 1).

The following spatio-temporal parameters were acquired (distances were determined by how the tip of the snout travelled in the direction of movement): speed (v in m/sec; motion of the snout tip), maximal height of bounds as determined by the distance of the tip of the snout to the ground (h_{aerial} in m), distance covered during bounds (s_{aerial} in m), distance covered while in contact with the ground (s_{ground} in m), total distance covered in one cycle (s_{total} in m), duration of the aerial phase (t_{aerial} in sec), duration of the contact phase (t_{ground} in sec), duration of the cycle (t_{total} in sec), and relative ground contact time (t_{rel}) as the percentage of t_{ground} to t_{total} . Box-and-whisker plots are used to visualize the distribution of parameter values (all subjects pooled) for the three experimental enclosures.

We used the non-parametric Mann-Whitney U test to test for significant differences of pooled data from all subjects for each gait parameter between the basic and experimentally modified enclosure dimensions (significance level was set to $p \leq 0.05$). Non-parametric tests are less requiring and more conservative than comparable parametric tests. Additionally, multivariate statistics were used to evaluate how changes of enclosure dimensions affected the overall locomotion pattern. To this end, PCA was performed in Python (version 3.6, Python Software Foundation, www.python.org). Importantly, there is a fundamental discrepancy between these pooled modes of uni- and multivariate analysis and the hierarchical character of the data: As is usual in motion analyses, we here analyze repeated measurements of four subject animals. Often, such data is dealt with in two alternative ways: (i) a pooled analysis of all observations and (ii) the analysis of only the averaged data. However, when pooling the data from these four subjects (i), the repeats are incorporated as pseudo-replication, disregarding the information about which animal they were acquired from. As we will show, this can lead to problems even in the case of balanced sample size across all individuals. In an analysis strategy that is based on

subjects' averages (ii), statistical power would be reduced (in our case, by reducing the number of observations to four). Instead of either of these unsatisfactory alternatives, we incorporated a hierarchical model structure as follows.

As an example of a hierarchical model, we constructed a hierarchical Bayesian GLM and used it to analyze the measurements. We modelled all parameters θ (i.e., locomotor parameters or principal components of those) by the formula

$$\theta_i = n_i + \sum_k m_{k,i} \cdot x_{k,i} + \varepsilon$$

Where k are the tested modified conditions ("long" and "high" enclosure, relative to "basic"), i are subject animals, $x_{k,i}$ are boolean vectors describing the respective settings under which each trial was recorded and ε denoted the residual. The model parameters n_i (subject intercepts) and $m_{k,i}$ (subject slopes) for each subject animal i are modelled as draws from a Normal distribution on population level:

$$\begin{aligned} n_i &\sim N(n, \sigma_n); m_{k,i} \sim N(m_k, \sigma_{m,k}) \\ n, m_k &\sim N(\bar{\mu}, \bar{\sigma}); \sigma_n, \sigma_{m,k} \sim \text{HalfCauchy}(\bar{\beta}) \end{aligned}$$

The population intercept value n indicates the population average of each parameter θ in the basic enclosure. The subject-specific intercepts n_i show subject-specific parameter values, containing all variance that might be attributed to potentially confounding individual differences of the studied subjects (like size, mass, age, sex, training level, level of anxiety during the experiments, etc.). The slopes $m_{k=long}$ and $m_{k=high}$ model the effect of the change from basic to the long and high enclosure, respectively, both on population level (m_k) and per subject ($m_{k,i}$). We were interested in effects on both levels. Each of the hyperparameters n , m_k (means) and σ (standard deviations) was initialized with weakly informative Normal ($\bar{\mu} = 0$, $\bar{\sigma} = 10$) and Half Cauchy ($\bar{\beta} = 5$) priors¹, where $\bar{\mu}$, $\bar{\sigma}$ and $\bar{\beta}$ are start values for the sampler. The model residual ε was set to be drawn from a Half Cauchy distribution (Gelman, 2006; Polson and Scott, 2012). We performed robust regressions by modelling our data as Student's t distributed with variable "degree of freedom" parameter (i.e., tail weight, herein abbreviated as η ; Gelman et al., 2013).

A MCMC sampling algorithm (Gilks et al., 1995) was employed using the PyMC3 package (version 3.1; Salvatier et al., 2016) in Python. We chose a NUTS sampler (Hoffman and Gelman, 2014) that was initialized by the ADVI algorithm (Kucukelbir et al., 2017), which is the default setting in PyMC3. Our

¹ In Bayesian terms, a "prior" is the probability density/distribution that initializes a model parameter, whereas the "posterior" is the distribution of values that the Bayesian formula returns after incorporation of the data and normalization.

model and code are based on an introductory blog post by Elbers and Wiecki (2014). The MCMC algorithm was run for twelve repetitions, 100,000 iterations each. Sampler convergence was verified *ex post* by calculating Gelman-Rubin statistic (\hat{R} close to one is indicative of convergence; Gelman and Rubin, 1992) and by visual inspection of the sample traces. Due to the low sample size and the complex shape of the posterior distribution, we experienced some cases in which the sampler eventually got static in a local area of the posterior distributions for brief sampling episodes (ca. 30% of the repetition runs). In such cases, the whole run was discarded and repeated. This occasional MCMC behavior is related to sampler step size adaptation and does not alter our results. We confirmed this by varying the sampler (the Slice sampler from Neal [2003] proved less susceptible), by changing the sampler initialization, and by testing a re-parametrization to decouple the population mean and standard deviation, which improves sampler coverage and helps sampler convergence (Betancourt and Girolami, 2015; Wiecki, 2017). All these modifications affected the probability of static episodes to occur, but never the overall position of the posterior distributions. Only every fifth sample value of the final 50,000 iterations for each of the twelve repetitions was kept, which means that in total 120,000 samples determined the overall shape of the posterior distributions. The 95% highest posterior density (HPD) intervals within the distribution of the simulated data were considered and it was checked whether or not these intervals included the value “0”. If not, the hypothesis of a difference of the analyzed spatio-temporal gait parameter caused by the experimentally altered enclosure dimensions compared to the basic enclosure was corroborated with a 95% probability (the “credible interval” in Bayesian terms; Gelman et al., 2013). If the value “0” is included in a HPD interval, the hypothesis was rejected.

3. Results

3.1 Comparison of gait parameters

Basic data per subject is reported in Table 2. To test whether the enclosure dimensions have any effect on the subject animals’ locomotor parameters, we compared pooled data between enclosures for each parameter. Despite the overall visible overlap of the different parameters’ data obtained in the three experimental conditions, several parameters of the squirrel’s bounding locomotion appeared to be affected by the experimental modification of the enclosure dimensions (Fig. 2). When aggregated over subject animals, the values obtained in the high enclosure for the parameters h_{aerial} , S_{ground} , t_{ground} , and t_{rel} turned out significantly different from the basic enclosure according to the non-parametric Mann-Whitney U test. Similarly, the values obtained in the long enclosure for the parameters v , S_{aerial} , S_{ground} , S_{total} , t_{ground} , and t_{rel} were significantly different from the basic enclosure.

However, the hierarchical GLM did not confirm these effects for any of the parameters (see Fig. 3; Table S1). In all cases the HPD interval of the population distribution of parameter changes (m_k) associated with the high or long enclosure contained the value “0” (indicated by red vertical lines in Fig.

3; Table S1). Notably, for most of the parameters, data of one or more subjects deviated from the overall outcome (see e.g., “pink” subject in Fig. 3). The model turns out to be robust against data bias introduced by some of the analyzed subjects in this study. Moreover, the hierarchical GLM provides evidence that locomotor parameters are robust against the considerable experimental variation of enclosure dimensions. Consequently, the hypothesis of a predictable parameter change due to the enclosure modification was rejected. Sampler convergence was confirmed by Gelman-Rubin statistics ($|\hat{R} - 1| < 4 * 10^{-3}$ for all parameters). The posterior value of the degree of freedom model parameter η of the Student’s t distribution was found to be 11.4 on average, ranging from 4.9 to 17.7 for all θ . These relatively high values indicate that the parameter value distributions’ shape has high similarity to a Normal distribution.

3.2 Comparison of overall gait pattern

To test whether any effects of enclosure dimensions are revealed in a multivariate analysis, we submitted our data to a PCA. PC1 corresponded to 54% of the variance of the entire dataset and the loading (visualized by the length of the bars in Fig. 4) indicated a strong positive correlation between most of the parameters except for t_{ground} , and s_{ground} along PC1 (Fig. 4; Table 3). Also, t_{rel} correlated negatively to the other parameters along PC1. PC2 corresponded to 25% of the total variance. Along PC2 the loading indicated a strong negative correlation between v and t_{ground} . The scatter plot showed that some of the variance along PC2 can be attributed to the data of the “pink” subject. PC3 accounted for 13% of the total variance with s_{ground} and t_{ground} showing the closest correlation. Comparison of PCs of the spatio-temporal gait parameters between enclosure conditions also showed significant Mann-Whitney U Test outcomes (Fig. 2). Both non-basic enclosures induced a significant change on PC1. The long and the basic enclosure also differed significantly with regards to PC2, whereas the change to the high enclosure affected PC3. Generally, both enclosure modifications lead to more variable data (as reflected in larger occupied areas with the respective convex hulls in Fig. 4).

Again, the hierarchical GLM did not confirm the outcomes of the significance tests (Fig. 5). The 95% HPD intervals of all of the comparisons between the basic enclosure and both of the experimentally modified enclosure dimensions included the value “0” and thus, the hypothesis of a predictable parameter change due to the enclosure dimensions was rejected in all cases (Table S1). All of the first three PCs (together corresponding to 92% of total variance, Table 3) in the hierarchical GLM proved robust against introduced data bias by the subjects that clearly deviated from the others (e.g., “pink” subject on PC2). Convergence tests were again positive ($|\hat{R} - 1| < 10^{-4}$ for all principal components) and the likelihood distribution’s tail weight low ($\eta=11.9$ on average).

In sum, using a model that is able to appropriately handle the effect of inter-individual variation, we found no evidence that uni- or multivariate descriptors of locomotor behavior are affected by the experimental modifications of the setup enclosure dimensions tested herein.

4. Discussion

In-vivo motion analyses of a limited number of subjects of non-model organisms are a widely used method to generate a quantitative basis for functional investigations of the vertebrate locomotor system. Data derived from such analyses is inherently hierarchical, a property that needs to be taken into account when drawing inferences on the level of statistical population (e.g., the species). We presented a study that exemplifies the difference on the statistical inference a hierarchical model can make compared to a widely used, non-hierarchical statistical test in the case of the hitherto understudied potential confounding influence of the setup enclosure dimensions. To provide the possibility to replicate our results and to adjust the hierarchical Bayesian GLM to other researchers' analyses, we provide our complete dataset and code (implemented in the freely available Python software environment) in the supplementary information. Whereas results from the commonly used hypothesis test indicated a variety of influences of enclosure dimensions on locomotion, the GLM did not provide evidence for any effect. The tests differ in two ways: the hierarchical structure and the Bayesian framework of the GLM. The Bayesian model framework used herein is only one of several possible approaches to account for the hierarchical nature. A non-hierarchical Bayesian GLM (as provided in the supplementary source code, but not shown here) yields results that are identical to those of the Mann-Whitney U test. Hence, the difference in the two reports is solely due to the hierarchical model structure. Alternatively to the here used Bayesian framework, General Linear Mixed Models are a commonly used technique that combines point estimates for "fixed" parameters and distribution estimates for "random" parameters. We herein chose a purely Bayesian approach, which is discussed to be more reliable on low sample size (Hox, 2010) and more flexible (Hong et al., 2013) to avoid some specific design restrictions (no negative variance estimates, Wakefield, 2013). Hierarchical models, like the hierarchical Bayesian GLM used here, can account for any confounding effects on the level of individual subjects (body mass, sex, age, etc.) and can elucidate a subject's specific locomotor response as well as the group's response to a given change in enclosure dimensions. Only when taking the diversity of locomotor adjustments into account, locomotor plasticity, which characterizes adaptive locomotion in small mammals, may be fully understood.

We analyzed a limited number of subjects using a hierarchical GLM. Our study demonstrated that the confounding effect of a single subject that deviates in its locomotor behavior is potentially bigger than the influence of the variation in experimental enclosure dimensions. The dependency on a limited number of subjects is a common restriction of comparative small mammal *in-vivo* motion analyses. This points to the potential non-uniformity of kinematic (and perhaps mechanical) adjustments of individuals

of the same species to specific functional demands posed by the support properties. Omitting particular trials or even data from particular subjects may obscure the variability of a species' locomotor adjustments (Andrada et al., 2015). In a hierarchical Bayesian model, each individual is modeled by itself, but in relation to the distribution on other levels. Due to this structure, hierarchical models yield appropriate results even in a situation of sample size difference across animals, i.e., there is no need to discard data for reasons of uniform sample size. In comparison to non-Bayesian models, Bayesian models can even be more accurate if sample size is low (Hox, 2010:273). Herein, we forced uniform sample size to improve the statistical power of the Mann Whitney U test. Still, the analyses with the hierarchical GLM did prove more robust than the Mann Whitney U test against inter-individual differences despite our limited population sample of just four study subjects. The reason for this is that specifying a Normal- or Student's t distribution of the model likelihood provides a model with the flexibility to adjust for how likely outliers are. This is a common property of models with non-Uniform likelihood distribution, but fostered herein by the use of Student's t distribution: subjects far from the mean on population level have relatively lower probability, making hierarchical models less prone to outliers (cf. Elbers and Wiecki, 2014). Herein, we left the distribution's tail mass variable and retrieved a relatively high posterior mean (average $\eta = 11.5$), indicating that the outliers we observe have only marginally higher probability than would be expected for a Normal distribution. Note that the limited number of subjects available for this study resulted in relatively wide posteriors (i.e., high standard deviation) on the higher hierarchical level, which makes it necessary to carefully interpret the distributions of the subjects (Figs. 3, 5). In general, the relative width of the distributions of model parameters on each hierarchical level is determined by the balance of sampling from the population and from subjects.

Only few studies used the dimensions of the runway enclosure such as height and width to constrain the locomotion on purpose. These studies are typically interested in how animals move in confined spaces like crevices or burrows and often focus on "invertebrate" taxa (e.g., Gravish et al., 2013; Jayaram et al. 2016). In their analysis of subterranean locomotion of domestic ferrets (*Mustela putorius*), Horner and Biknevičius (2010) simulated the burrows of typical prey using height-constrained acrylic glass tunnels and forced a crouching posture with reduced maximum back height and hip height by 40% and 25%, respectively. Surprisingly, these authors found only a few significant differences to unconstrained over ground locomotion in *M. putorius* and attributed their findings to the species' adaptation to underground hunting. In a later study, Horner et al. (2016) studied the energetic costs of height constrained underground locomotion in comparison to unconstrained over ground locomotion in *M. putorius* and degus (*Octodon degu*) and found increased locomotor costs in the confined environment likely driven by postural changes. In contrast to these dramatic examples of a constraining runway enclosure, we varied the enclosure dimensions within what we deemed "non-constraining" dimensions. Potentially, our tested dimensions could have therefore been chosen intuitively by researchers that are interested in general locomotor characteristics of this species. Nevertheless, we conducted considerable

changes as we varied enclosure height by 100% and enclosure length by 50%. All three of our experimental enclosure dimensions allowed for the animals to have at least three to four locomotor cycles with extended aerial phases and allowed for a maximum jumping height of several limb lengths. At no time did we observe the animals to make contact with the side walls or any other parts of the runway enclosure during locomotor bouts in any of the analyzed trials. This indicates that, despite the transparency of the enclosure, the animals were aware of it. We initially anticipated spatio-temporal parameters to be affected by our experimental changes of the enclosure dimensions. However, in light of the limited effect of a dramatically confined runway on the spatio-temporal parameters of *M. putorius* and *O. degu* (Horner and Biknevicius, 2010), it might not surprise that we found no statistical support for predictable differences due to changes of enclosure dimensions in the spatio-temporal locomotor parameters of *T. swinhoei* using the hierarchical model. Importantly, this was only revealed by using the GLM, which takes the hierarchical nature of the data into account, but not using a widely used non-hierarchical statistical test.

Several other studies found changes of a suite of spatio-temporal parameters with an increase of speed during asymmetric gaits. Oftentimes changes of speed and correlated parameters of small mammals were observed when support characteristics were experimentally altered to test for adjustments to ecologically relevant differences of support properties (e.g., Lammers and Biknevicius, 2004; Nyakatura et al., 2008; Stevens, 2008; Young et al., 2009; Schmidt and Fischer, 2011; Schmidt, 2014; Shapiro et al., 2016; Karantanis et al., 2015; 2016; 2017). A clear dependency of locomotor speed on the dimensions of the three non-constraining setup enclosures we tested herein would have complicated comparisons between data of species derived from different setups with or without enclosures or with enclosures of different dimensions. However, the hierarchical GLM revealed locomotor speed of a highly agile mammal species to not be influenced by the considerable differences in setup enclosure dimensions we compared. Therefore, differences in locomotor parameters found in other species indeed likely reflect mainly the ecologically relevant alterations of support characteristics in the studies cited above. In regard of enclosure dimensions, the statistical evidence for most subjects' indifference to the considerable experimental variation of the dimensions in our study justifies the common practice of intuitively selecting dimensions that the researcher deems as "non-constraining". In order to facilitate comparison of data across studies and species, however, we nevertheless encourage researchers to report the dimensions not only of the trackways and/or the ecologically relevant experimental manipulations of the support properties, but also the dimensions of the runway enclosure.

We conclude that the Bayesian GLM used herein (and probably other hierarchical models) is suitable if measurements are repeatedly acquired from a limited number of subjects. This is due to the hierarchical structure of the model exploiting all the available information. Since Bayesian statistics work on density distributions instead of point estimates, they yield the statistical distribution of possible actual values of the parameter, given a careful (i.e., a non-biasing) specification of the prior (cf. Gelman

et al., 2013). When sample size is low and little data is available—a typical restriction in the comparative *in-vivo* motion analysis of non-model organisms—density distributions are generally wider, reflecting the high uncertainty after only few measurements. However, even wide distributions remain reliable in a sense that they depict the parameter range that is realistically given by the data at hand. Because our case study and other *in-vivo* motion analyses are inherently hierarchical, information is best conveyed by hierarchical models. While the proposed Bayesian framework is well suited for the problem addressed here, it is immensely flexible and not restricted to linear models. Finally, models can be set up to quantify the deviance of model uncertainty from normality and the influence of outliers (herein η parameter). To guide others in the application of this type of analysis, we provide our full dataset and all code necessary for the model calculation in the supplementary materials.

Data availability

All code and data are provided in the open Dryad data base (<https://doi.org/10.5061/dryad.10rn5>; Mielke et al., 2018) and have been assigned a digital object identifier (doi:10.5061/dryad.10rn5; ESM files “Mielke et al_ESM_data.” and “Mielke et al_ESM_code”).

Author contributions

FM, JW, and JAN conceived the study, VS and JW conducted the experiments, FM performed statistical analyses, FM and JAN wrote the manuscript. All authors contributed to the interpretation of the results and revised and approved the manuscript.

Acknowledgements

The authors would like to thank Lorenz Husterer for making the animals available for this study. We thank Thomas Wiecki, Adrian Seyboldt and Junpeng Lao for support and feedback on our initial model (discourse.pymc.io). The authors further thank Maja Mielke, Eli Amson, and two anonymous reviewers

for helpful critique of an earlier version of the manuscript. The study was funded by the German Research Council (DFG EXC 1027). The authors declare no competing interests.

References

- Albert, J.H., 1988. Computational methods using a Bayesian hierarchical generalized linear model; *J. Am. Stat. Assoc.*, 83(404):1037-1044.
- Amson, E., Nyakatura, J.A., 2017. The postcranial musculoskeletal system of xenarthrans: Insights from over two centuries of research and future directions. *J. Mamm. Evol.* online first, doi:10.1007/s10914-017-9408-7.
- Andrada, E., Haase, D., Sutedja, Y., Nyakatura, J. A., Kilbourne, B. M., Denzler, J., Fischer, M.S., Blickhan, R., 2015. Mixed gaits in small avian terrestrial locomotion. *Sci. Rep.*, 5.
- Ashley-Ross, M. A., Bechtel, B. F., 2004. Kinematics of the transition between aquatic and terrestrial locomotion in the newt *Taricha torosa*. *J. Exp. Biol.*, 207(3), 461-474.
- Autumn, K., Hsieh, S.T., Dudek, D.M., Chen, J., Chitaphan, C., Full, R.J., 2006. Dynamics of geckos running vertically. *J. Exp. Biol.*, 209(2), 260-272.
- Betancourt, M., Girolami, M., 2015. Hamiltonian Monte Carlo for hierarchical models. *Current trends in Bayesian Meth. with applications*, 79, 30.
- Biknevicius, A.R., Reilly, S.M., McElroy, E.J., Bennett, M.B., 2013. Symmetrical gaits and center of mass mechanics in small-bodied, primitive mammals. *Zoology*, 116(1), 67-74.
- Birn-Jeffery, A.V., Daley, M.A., 2012. Birds achieve high robustness in uneven terrain through active control of landing conditions. *J. Exp. Biol.*, 215(12), 2117-2127.
- Birn-Jeffery, A.V., Higham, T.E., 2014. Geckos significantly alter foot orientation to facilitate adhesion during downhill locomotion. *Biol. Lett.*, 10(10), 20140456.
- Bundle, M.W., Dial, K.P., 2003. Mechanics of wing-assisted incline running (WAIR). *J. Exp. Biol.*, 206(24), 4553-4564.
- Chadwell, B.A., Young, J.W., 2015. Angular momentum and arboreal stability in common marmosets (*Callithrix jacchus*). *Am. J. Phys. Anthropol.*, 156(4), 565-576.
- Cosans, C.E., Frampton, M., 2015. History of Comparative Anatomy. eLS. 1–8.
- Daley, M.A., Biewener, A.A., 2003. Muscle force-length dynamics during level versus incline locomotion: a comparison of in vivo performance of two guinea fowl ankle extensors. *J. Exp. Biol.*, 206(17), 2941-2958.
- Dempster, A.P., Selwyn, M.R., Weeks, B.J., 1983: Combining Historical and Randomized Controls for Assessing Trends in Proportions; *J. Ame. Stat. Assoc.*, 78(382):221-227.
- Duhamel, A., Bourriez, J.L., Devos, P., Krystkowiak, P., Destée, A., Derambure, P., Defebvre, L., 2004. Statistical tools for clinical gait analysis. *Gait and Posture*, 20, 204-212.
- Elbers, D., Wiecki, T., 2014: The Best Of Both Worlds: Hierarchical Linear Regression in PyMC3; Blog "While My MCMC Gently Samples", <http://twiecki.github.io/blog/2014/03/17/bayesian-glms-3> (accessed 2017/06/02).

- Foster, K.L., Higham, T.E., 2012. How forelimb and hindlimb function changes with incline and perch diameter in the green anole, *Anolis carolinensis*. *J. Exp. Biol.*, 215(13), 2288-2300.
- Franz, T.M., Demes, B., Carlson, K.J. 2005. Gait mechanics of lemurid primates on terrestrial and arboreal substrates. *J. Hum. Evol.*, 48(2), 199-217.
- Gelman, A., 2006. Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper); *Bayesian Analysis*, 1(3):515-534.
- Gelman, A., Rubin, D.B. (1992). Inference from iterative simulation using multiple sequences. *Stat. Sci.*, 457-472.
- Gelman, A., Carlin, J.B., Stern, H.S., Dunson, D.B., Vehtari, A., Rubin, D.B., 2013, *Bayesian Data Analysis*; 3 ed. Texts in Statistical Science Series, Chapman & Hall/CRC press, Boca Raton.
- Gilks, W.R., Richardson, S., Spiegelhalter, D., 1995, *Markov Chain Monte Carlo In Practice*; Chapman & Hall/CRC press, London.
- Gillis, G. B., 1998. Environmental effects on undulatory locomotion in the American eel *Anguilla rostrata*: kinematics in water and on land. *J. Exp. Biol.*, 201(7), 949-961.
- Good, I.J., 1980, Some history of the hierarchical Bayesian methodology; *Trabajos de Estadística Y de Investigación Operativa*, 31(1):489.
- Gravish, N., Monaenkova, D., Goodisman, M.A., Goldman, D.I., 2013. Climbing, falling, and jamming during ant locomotion in confined environments. *Proc. Nat. Acad. Sci. U.S.A.*, 110(24), 9746-9751.
- Herrel, A., Perrenoud, M., Decamps, T., Abdala, V., Manzano, A., Pouydebat, E., 2013. The effect of substrate diameter and incline on locomotion in an arboreal frog. *J. Exp. Biol.*, 216(19), 3599-3605.
- Hesse, B., Nyakatura, J.A., Fischer, M.S., Schmidt, M., 2015. Adjustments of limb mechanics in cotton-top tamarins to moderate and steep support orientations: significance for the understanding of early primate evolution. *J. Mamm. Evol.*, 22(3), 435-450.
- Hildebrand, M., 1977. Analysis of Asymmetrical Gaits. *J. Mammal.* 58:131-156.
- Hildebrand, M., 1989. The quadrupedal gaits of vertebrates. *BioScience*, 39(11), 766.
- Hoffman, M.D., Gelman, A., 2014. The No-U-turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo.; *J. Machine Learning Research*, 15(1):1593-1623.
- Hong, H., Carlin, B.P., Shamliyan, T.A., Wyman, J.F., Ramakrishnan, R., Sainfort, F., Kane, R.L., 2013. Comparing bayesian and frequentist approaches for multiple outcome mixed treatment comparisons. *Medical Decision Making*, 33(5), 702-714.
- Horner, A.M., Biknevicius, A.R., 2010. A comparison of epigeal and subterranean locomotion in the domestic ferret (*Mustela putorius furo*: Mustelidae: Carnivora). *Zoology*, 113(3), 189-197.
- Horner, A.M., Hanna, J.B., Biknevicius, A.R., 2016. Crouching to fit in: the energetic cost of locomotion in tunnels. *J. Exp. Biol.*, 219(21), 3420-3427.
- Hox, J.J., 2010. *Multilevel Analysis: Techniques and Applications*, 2nd ed. Quantitative methodology series, Routledge Academic, New York.
- Iknayan, K.J., Tingley, M.W., Furnas, B.J., Beissinger, S.R., 2014. Detecting diversity: emerging methods to estimate species diversity; *Trends Ecol. Evol.*, 29(2):97 - 106.
- Jayaram, K., Full, R.J., 2016. Cockroaches traverse crevices, crawl rapidly in confined spaces, and inspire a soft, legged robot. *Proc. Nat. Acad. Sci. U.S.A.*, 113(8), E950-E957.
- Jayne, B.C., Ellis, R.V., 1998. How inclines affect the escape behaviour of a dune-dwelling lizard, *Uma scoparia*. *Anim. Behav.*, 55(5), 1115-1130.
- Karantanis, N.E., Rychlik, L., Herrel, A., Youlatos, D., 2016. Vertical locomotion in *Micromys minutus* (Rodentia: Muridae): insights into the evolution of eutherian climbing. *J. Mamm. Evol.*, 1-13.

- Karantanis, N.E., Rychlik, L., Herrel, A., Youlatos, D., 2017. Arboreal gaits in three sympatric rodents *Apodemus agrarius*, *Apodemus flavicollis* (Rodentia, Muridae) and *Myodes glareolus* (Rodentia, Cricetidae). *Mamm. Biol.*, 83, 51-63.
- Karantanis, N.E., Youlatos, D., Rychlik, L., 2015. Diagonal gaits in the feathertail glider *Acrobates pygmaeus* (Acrobatidae, Diprotodontia): insights for the evolution of primate quadrupedalism. *Journal of human evolution*, 86, 43-54.
- Kramer, D.L., McLaughlin, R.L., 2001. The behavioral ecology of intermittent locomotion. *Am. Zool.* 41(2), 137-153.
- Kucukelbir, A., Tran, D., Ranganath, R., Gelman, A., Blei, D.M., 2017. Automatic Differentiation Variational Inference; *Journal of Machine Learning Research*, 18(14):1-45.
- Lammers, A.R., Biknevicius, A.R., 2004. The biodynamics of arboreal locomotion: the effects of substrate diameter on locomotor kinetics in the gray short-tailed opossum (*Monodelphis domestica*). *J. Exp. Biol.*, 207(24), 4325-4336.
- Lammers, A.R., Earls, K.D., Biknevicius, A.R., 2006. Locomotor kinetics and kinematics on inclines and declines in the gray short-tailed opossum *Monodelphis domestica*. *J. Exp. Biol.*, 209(20), 4154-4166.
- Lee, M.D., 2011, How cognitive modelling can benefit from hierarchical Bayesian models. *J. Math. Psych.*, 55(1):1 - 7.
- Lindley, D.V., Smith, A.F.M., 1972, Bayes estimates for the linear model. *J. Roy. Stat. Soc.. Series B (Methodological)*, 34(1):1-41.
- Marey, E.J., 1886. *La machine animale: locomotion terrestre et aérienne* (Vol. 3). Germer Baillière.
- Mielke, F., Schunke, V., Wölfer, J., Nyakatura, J.A., 2018. Data from: Motion analysis of non-model organisms using a hierarchical model: Influence of setup enclosure size on gait parameters of Swinhoe's striped squirrels as a test case. *Dryad Digital Repository*. doi:10.5061/dryad.10rn5.
- Muybridge, E., 1878. The science of the horse's motions. *Sci. Am.*, 39, 241.
- Neal, R., 2003. Slice sampling. *The Analysis of statistics*, 31(3), 705-741.
- Nyakatura, J.A., Andrada, E., Grimm, N., Weise, H., Fischer, M.S., 2012. Kinematics and center of mass mechanics during terrestrial locomotion in northern lapwings (*Vanellus vanellus*, Charadriiformes). *J. Exp. Zool. A: Ecol. Gen. Physiol.*, 317(9), 580-594.
- Nyakatura, J.A., Fischer, M.S., Schmidt, M., 2008. Gait parameter adjustments of cotton-top tamarins (*Saguinus oedipus*, Callitrichidae) to locomotion on inclined arboreal substrates. *Am. J. Phys. Anthropol.*, 135(1), 13-26.
- Polson, N.G., Scott, J.G., 2012. On the half-cauchy prior for a global scale parameter. *Bayesian Analysis*, 7(4):887--902.
- Rollinson, J., Martin, R.D., 1981. Comparative aspects of primate locomotion, with special reference to arboreal cercopithecines. In. *Symp. Zool. Soc. Lond.*, 48, 377-427.
- Rouder, J.N., Lu, J., 2005. An introduction to Bayesian hierarchical models with an application in the theory of signal detection; *Psychon. Bull. Rev.*, 12(4), 573-604.
- Salvatier, J., Wiecki, T.V., Fonnesbeck, C., 2016, Probabilistic programming in Python using PyMC3; *PeerJ Computer Science*, 2:e55.
- Schmidt, A., 2014. Locomotion in degus on terrestrial substrates varying in orientation—implications for biomechanical constraints and gait selection. *Zoology*, 117(2), 146-159.
- Schmidt, A., Fischer, M.S. 2010. Arboreal locomotion in rats—the challenge of maintaining stability. *J. Exp. Biol.*, 213(21), 3615-3624.

- Schmidt, A., Fischer, M.S., 2011. The kinematic consequences of locomotion on sloped arboreal substrates in a generalized (*Rattus norvegicus*) and a specialized (*Sciurus vulgaris*) rodent. *J. Exp. Biol.*, 214(15), 2544-2559.
- Shapiro, L.J., Kemp, A.D., Young, J.W., 2016. Effects of substrate size and orientation on quadrupedal gait kinematics in mouse lemurs (*Microcebus murinus*). *J. Exp. Zool. A: Ecol. Gen. Physiol.*, 325(5), 329-343.
- Shapiro, L.J., Young, J.W., 2010. Is primate-like quadrupedalism necessary for fine-branch locomotion? A test using sugar gliders (*Petaurus breviceps*). *J. Hum. Evol.*, 58(4), 309-319.
- Stevens, N.J., 2008. The effect of branch diameter on primate gait sequence pattern. *Am. J. Primatol.*, 70(4), 356-362.
- Tickle, P.G., Nudds, R.L., Codd, J.R., 2012. Barnacle geese achieve significant energetic savings by changing posture. *PLoS One*, 7(10), e46950.
- Vanden Hole, C., Goyens, J., Prims, S., Franssen, E., Hernando, M.A., Van Cruchten, S., Aerts, P., Van Ginneken, C., 2017. How innate is locomotion in precocial animals? A study on the early development of spatio-temporal gait variables and gait symmetry in piglets. *J. Exp. Biol.*, 220, 2706-2716.
- Wainwright, P.C., Reilly, S.M., 1994. *Ecological Morphology: Integrative Organismal Biology*. University of Chicago Press, Chicago.
- Wakefield, J., 2013. *Bayesian and Frequentist Regression Methods*. Springer Series in Statistics, Springer, New York.
- Wiecki, T., 2014: Why hierarchical models are awesome, tricky, and Bayesian; Blog "While My MCMC Gently Samples", <http://twiecki.github.io/blog/2017/02/08/bayesian-hierarchical-non-centered> (accessed 2017/08/10)
- Wikle, C.K., 2003. Hierarchical Bayesian models for predicting the spread of ecological processes; *Ecology*, 84(6):1382-1394.
- Youlatos, D., Karantanis, N.E., Byron, C.D., Panyutina, A., 2015. Pedal grasping in an arboreal rodent relates to above-branch behavior on slender substrates. *J. Zool., Lond.*, 296(4), 239-248.

Figure Captions

Fig. 1: A locomotor bout of *Tamiops swinhoei*. From left to right: Forelimb touch-down, mid-contact, hindlimb lift-off, mid-aerial, subsequent forelimb touch-down.

(Fig.1: Intended for double column width)

Fig. 2: Beeswarm box-and-whisker plots for the different spatio-temporal parameters and the first three PCs of the high (length x height x width: 200 cm x 40 cm x 20 cm), basic (200 cm x 20 cm x 20 cm), and long (300 cm x 20 cm x 20 cm) enclosure conditions, respectively. Trials of all subjects were pooled for each parameter for the box plots, but offset. Subjects are color coded as follows: subject 1, green; subject 2, pink, subject 3, yellow; subject 4, purple (cf. Table 1). Grey bars represent significant (at the $p = 0.05$ level) differences as determined by the Mann-Whitney U tests in comparisons of the basic enclosure condition with each of the two modified enclosure conditions (i.e., “basic to long” and “basic to high”).

(Fig. 2: Intended for double column width)

Fig. 3: Hierarchical generalized linear model (GLM) results for the spatio-temporal gait parameters for comparison of the two experimentally modified conditions with the basic condition (expressed as the difference Δ to the basic condition). Relative changes of locomotor parameters (slope $m_{k,i}$ divided by intercept n_i for each subject, in %) are shown on the x-axes. The effect of enclosure dimension change from "basic to high" is shown in the left column, whereas "basic to long" is shown in the right column. The red vertical line indicates 0% change. Black and colored (color code cf. Table 1, Fig. 2) distribution curves show the posterior probability density of the parameter value on population and subject level, respectively, with their most likely value marked (vertical lines). Their height is arbitrarily standardized here (the area each distribution curve covers is 1 per definition) to emphasize the hierarchical relation of the parameter on population level and the subjects' specific results. Each curve is based on 120,000 MCMC sampling iterations (12 repetitions à 100,000 samples, burn-in 50,000 and thinning 5). In order to reject the null hypothesis of a predictable change of parameter values due to the experimental shift of the enclosure dimension the red vertical line should not be included in the 95% HPD interval (dashed lines separated by the arrow).

(Fig. 3: Intended for double column width)

Fig. 4: Scatter plots of all trials coded for subjects (color code cf. Table 1, Fig. 2) and experimental enclosures including loadings of analyzed parameters on the first three PCs. Convex hulls depict occupied parameter space for the three enclosure conditions (see inset legend).

(Fig. 4: Intended for 1.5 column width)

Fig. 5: Hierarchical GLM results for the first three PCs for comparison of the two experimentally modified conditions with the basic condition. Absolute changes of PCs are shown in rows. Otherwise, the display is analogous to that of the univariate parameters (Fig. 4.).

(Fig. 5: Intended for double column width)

ACCEPTED MANUSCRIPT

Fig 1



ACCEPTED MANUSCRIPT

Fig 2

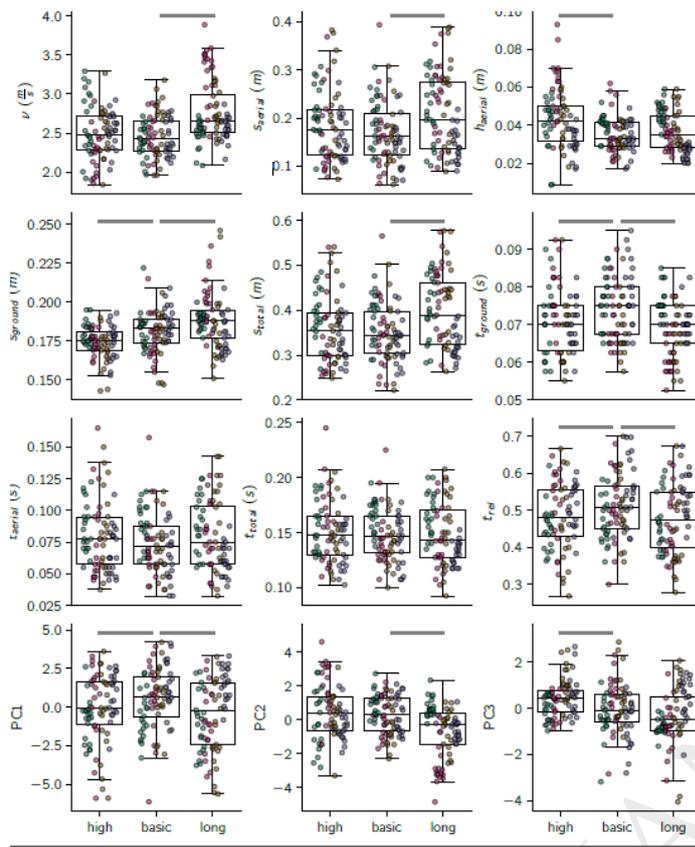


Fig 3

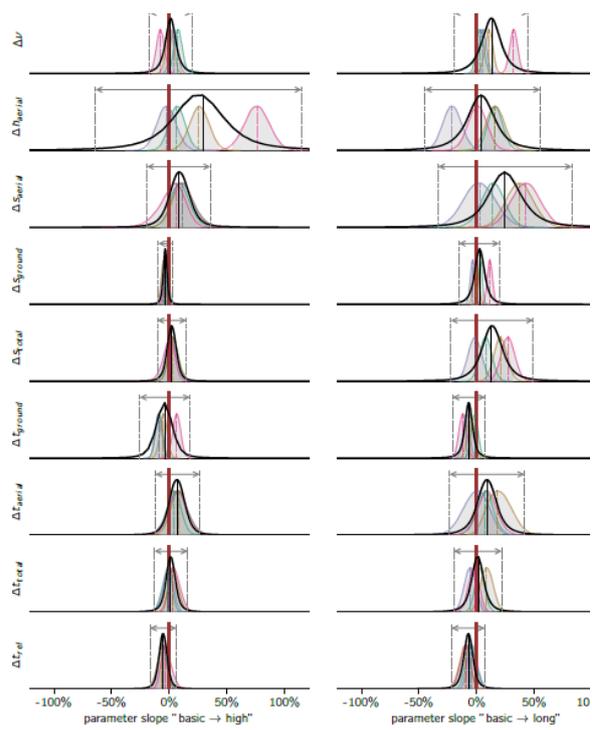


Fig 4

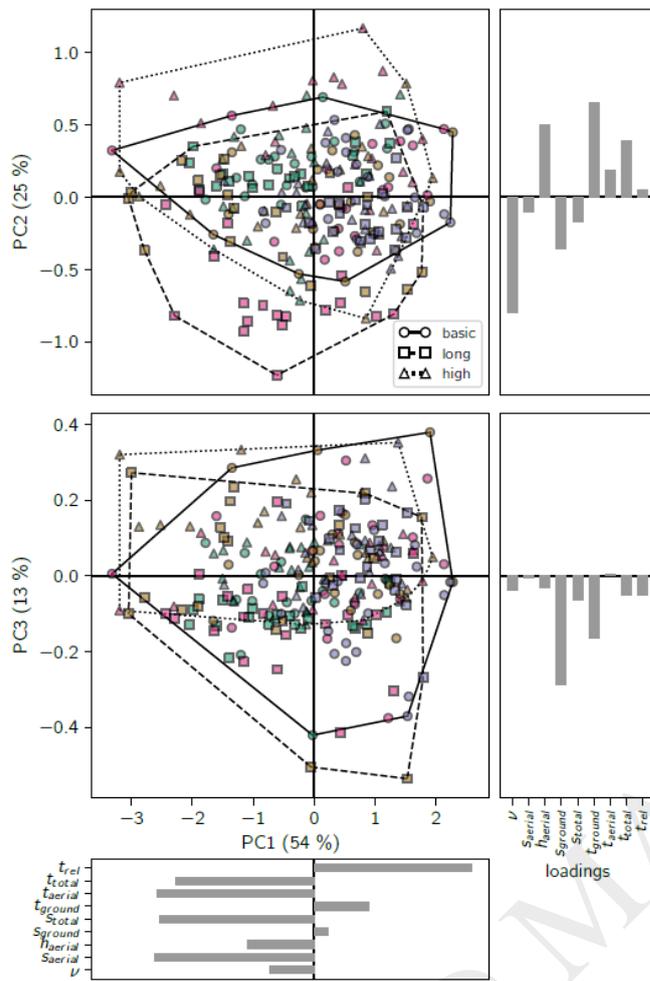
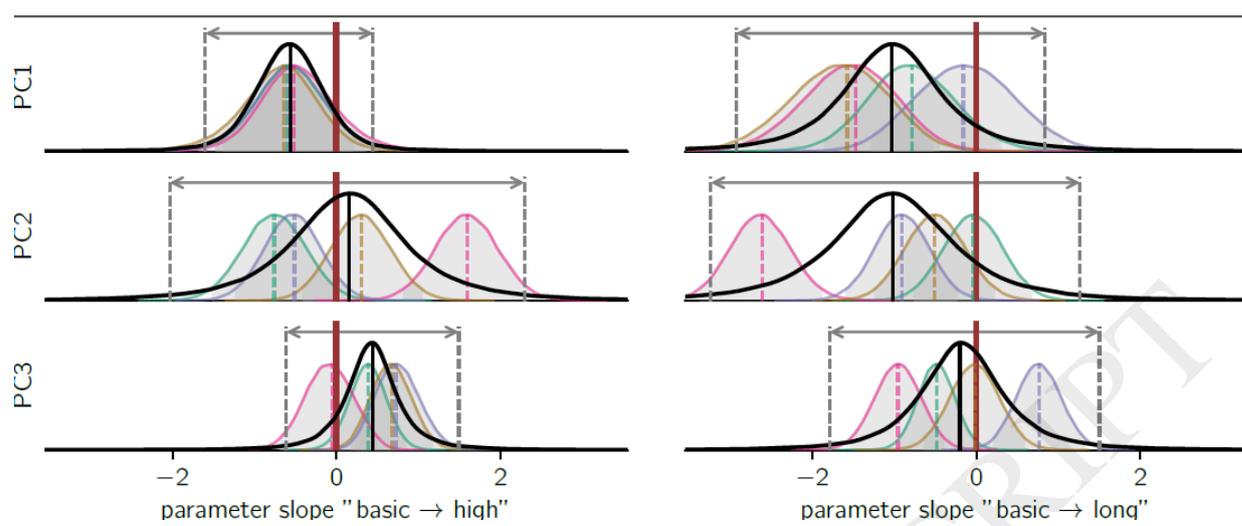


Fig 5



Tables

Table 1: Study subjects.

Subject	*color	Year of birth	Weight in g	Crown-rump-length in cm	No. of trials basic enclosure	No. of trials high enclosure	No. of trials long enclosure
1	green	2014	106	13.77	20	20	20
2	pink	2016	85	12.67	20	18	20
3	yellow	2016	76	13.37	20	20	20
4	purple	2014	88	12.80	20	20	20

*: Color corresponds to the labelling of the study subjects in the figures.

Table 2: Basic statistics of the analyzed trials per subject. Data reported as mean +/- standard deviation.

subject	Speed (m/sec)	h_{aerial} (m)	s_{ground} (m)	s_{aerial} (m)	s_{total} (m)	t_{ground} (sec)	t_{aerial} (sec)	t_{total} (sec)	t_{ret}
Basic enclosure									
green	2.38 (\pm 0.21)	0.04 (\pm 0.01)	0.18 (\pm 0.01)	0.20 (\pm 0.05)	0.38 (\pm 0.05)	0.08 (\pm 0.01)	0.08 (\pm 0.02)	0.17 (\pm 0.01)	0.47 (\pm 0.06)
pink	2.47 (\pm 0.26)	0.03 (\pm 0.01)	0.18 (\pm 0.01)	0.17 (\pm 0.08)	0.35 (\pm 0.08)	0.07 (\pm 0.01)	0.07 (\pm 0.03)	0.14 (\pm 0.03)	0.51 (\pm 0.10)
yellow	2.56 (\pm 0.35)	0.03 (\pm 0.01)	0.18 (\pm 0.02)	0.17 (\pm 0.06)	0.34 (\pm 0.07)	0.07 (\pm 0.01)	0.07 (\pm 0.02)	0.14 (\pm 0.02)	0.51 (\pm 0.09)
purple	2.42 (\pm 0.22)	0.03 (\pm 0.01)	0.19 (\pm 0.01)	0.14 (\pm 0.05)	0.33 (\pm 0.05)	0.08 (\pm 0.01)	0.06 (\pm 0.02)	0.14 (\pm 0.02)	0.56 (\pm 0.08)
High enclosure									
green	2.63 (\pm 0.41)	0.04 (\pm 0.01)	0.18 (\pm 0.01)	0.21 (\pm 0.06)	0.39 (\pm 0.06)	0.07 (\pm 0.01)	0.09 (\pm 0.02)	0.16 (\pm 0.02)	0.45 (\pm 0.06)
pink	2.25 (\pm 0.28)	0.06 (\pm 0.02)	0.17 (\pm 0.01)	0.17 (\pm 0.08)	0.34 (\pm 0.08)	0.08 (\pm 0.01)	0.08 (\pm 0.03)	0.16 (\pm 0.04)	0.51 (\pm 0.10)
yellow	2.53 (\pm 0.31)	0.04 (\pm 0.01)	0.17 (\pm 0.01)	0.20 (\pm 0.09)	0.37 (\pm 0.09)	0.07 (\pm 0.03)	0.08 (\pm 0.03)	0.15 (\pm 0.03)	0.46 (\pm 0.11)
purple	2.54 (\pm 0.23)	0.03 (\pm 0.01)	0.18 (\pm 0.01)	0.16 (\pm 0.05)	0.34 (\pm 0.05)	0.07 (\pm 0.01)	0.07 (\pm 0.02)	0.14 (\pm 0.02)	0.51 (\pm 0.06)
Long enclosure									
green	2.49 (\pm 0.17)	0.05 (\pm 0.01)	0.19 (\pm 0.01)	0.22 (\pm 0.07)	0.40 (\pm 0.07)	0.08 (\pm 0.00)	0.09 (\pm 0.02)	0.17 (\pm 0.02)	0.46 (\pm 0.08)
pink	3.24 (\pm 0.38)	0.03 (\pm 0.01)	0.20 (\pm 0.02)	0.24 (\pm 0.08)	0.44 (\pm 0.08)	0.06 (\pm 0.01)	0.08 (\pm 0.03)	0.14 (\pm 0.03)	0.46 (\pm 0.09)
yellow	2.77 (\pm 0.26)	0.04 (\pm 0.01)	0.18 (\pm 0.02)	0.24 (\pm 0.09)	0.42 (\pm 0.09)	0.07 (\pm 0.01)	0.09 (\pm 0.03)	0.16 (\pm 0.03)	0.44 (\pm 0.11)
purple	2.56 (\pm 0.16)	0.03 (\pm 0.00)	0.18 (\pm 0.01)	0.14 (\pm 0.03)	0.32 (\pm 0.03)	0.07 (\pm 0.00)	0.06 (\pm 0.01)	0.13 (\pm 0.01)	0.54 (\pm 0.06)

Table 3: Eigenvalues and % of variance per principal component (PC).

no. of PC	eigenvalue	% of variance
1	4.87	54.1
2	2.29	25.4
3	1.19	13.2
4	0.60	6.6
5	0.04	0.4
6	0.02	0.2
7	0.01	0.1
8	< 0.01	< 0.01
9	< 0.01	< 0.01