Journal of Equine Veterinary Science 115 (2022) 104024

Contents lists available at ScienceDirect

Journal of Equine Veterinary Science

journal homepage: www.j-evs.com

Determining Objective Parameters to Assess Gait Quality in Franches-Montagnes Horses for Ground Coverage and Over-Tracking - Part 1: At Walk

Annik Imogen Gmel[a,b,*], Eyrún Halla Haraldsdóttir[b], Filipe M. Serra Bragança[c], Antonio M. Cruz[d], Markus Neuditschko[a], Michael A. Weishaupt[b]

[a] Agroscope, Animal GenoPhenomics, Posieux, Switzerland
[b] Equine Department, Vetsuisse Faculty University of Zurich, Zurich, Switzerland
[c] Department of Clinical Sciences, Faculty of Veterinary Medicine, Utrecht University, Utrecht, The Netherlands
[d] Klinik für Pferdechirurgie und Orthopädie, Justus-Liebig Universität Giessen, Giessen, Germany

ARTICLE INFO

Article history:
Received 19 February 2022
Received in revised form 23 May 2022
Accepted 25 May 2022
Available online 29 May 2022

Keywords:
Equine
Walk
Speed
Stride length
Over-tracking distance
Breeding

ABSTRACT

Ground coverage and over-tracking are two gait quality traits describing the forward movement of the front respectively the hind limbs in relation to stride length and over-tracking distance. To investigate the complex interplay of different movement patterns in ground coverage and over-tracking, limb and body kinematics of 24 Franches-Montagnes (FM) stallions were measured with 3D optical motion capture (OMC) on a treadmill during an incremental speed test at the walk (1.4–2.0 m/s). The significance and amount of explained variance of kinematic parameters on stride length and over-tracking distance were estimated using linear mixed-effect models, with speed and horse as random effects. Two separate models were tested: a full model with all parameters measurable by OMC, and a reduced model with a subset of parameters also measurable with inertial measurement units (IMUs). The kinematic parameters were correlated to the subjective scores from six breeding experts to interpret their external validity. The parameter for ground coverage at the walk, explaining most of the variance in stride length, were the maximal forelimb retraction angle (11%) measured with OMC, and the range of pelvis pitch (10%) if measuring with IMUs. The latter was also the most relevant for quantifying over-tracking, explaining 24% to 33% of the variance in the over-tracking distance. The scores from most breeding experts were significantly correlated ($r \geq |0.41|$) with the fore- and hind limb protraction angles, which reflect the textual definition of ground coverage and over-tracking. Both gait quality traits can be objectively quantified using either OMC or IMUs.

1. Introduction

Gait quality, the way horses move according to functional and aesthetic principles, is expected to be predictive for future performance, which is one of the major breeding goals of European sport horses [1]. However, selecting horses based on gait quality traits such as animation or harmony is difficult, as gait quality traits are subject to personal interpretations by breeding experts. For example, the subjective assessment of equine movement patterns of 24 Franches-Montagnes (FM) stallions showed poor inter-rater reliabilities, suggesting differences in interpretation of the scored traits, and those at walk in particular [2]. Similarly, the agreement of scores from dressage judges during competitions has been considered unsatisfactory in several studies [3–6]. Likewise, it has been demonstrated that subjective lameness assessments also showed poor to moderate reliability [7–9] or lacked consistency with objective kinematic measurements [10].

Conflict of interest statement: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Ethical statement: This study was conducted under the animal experiment permit number VD 3164 complying to Swiss Federal Legislation. No animal was harmed during the experiment.

[*] Corresponding author at: Annik Imogen Gmel, PhD, Agroscope, Animal GenoPhenomics, Route de la Tioleyre 4, 1725 Posieux, Switzerland.

E-mail address: annik.gmel@agroscope.admin.ch (A.I. Gmel).

https://doi.org/10.1016/j.jevs.2022.104024

Despite the potential low reliability of gait assessments by human experts, the relevant kinematic and kinetic parameters for gait quality were identified by comparing elite and leisure horses [11], or by correlating measurements to a performance score attributed by a breeding expert [12,13] or dressage judge [14,15]. At walk, among the best-associated parameters with performance were symmetry of force vectors during braking and propulsion between contralateral limbs, stride rate and the relative step duration measured with accelerometers (Equimétrix) [14]. In a breeding context, some of these parameters – for example stride length (SL), stride duration and hind limb stance duration – have shown medium to high heritabilities in Pura Raza Espanol horses measured on the treadmill [16]. Considering the potentially low reliability of subjective scoring, the current advances in kinematic technologies for gait assessment allow for quantitative and a more detailed observation of equine locomotion patterns with optical motion capture (OMC). However, the costs and complexity of setting up and using an OMC system should not be underestimated (constant lighting, high-speed cameras, orientation of the horse in space, etc.), and often preclude measuring a large number of horses. The long-term aim is therefore to determine kinematic parameters which enable an objective quantification of gait quality traits of several hundreds of horses in the field using inertial measurement units (IMUs). This study aimed at identifying kinematic parameters for two gait quality traits, namely, ground coverage and hind limb engagement in the Franches-Montagnes (FM) breed, based on their textual definition.

In the FM breed, a light draught horse native to Switzerland, the walk should be "ground covering, supple and well cadenced" as defined in their breeding goals [17]. The term ground coverage (also scope or amplitude) can foremost be interpreted as a long stride: the longer the stride, the more ground is covered in one stride [2]. Additionally, ground coverage is also associated with the visual impression of the horse's forelimbs reaching both upward and onwards during forward movement [2]. Previous eye-tracking studies in dressage judges have shown an increase in the number and duration of fixations on movements of the forehand in comparison to the hindquarters [18,19]. As ground coverage is associated with the forelimb movement, it can be hypothesised that this trait has a major impact on gait quality. However, both the forelimb movement and stride length (SL) need to be considered simultaneously, as an elevated movement of the forelimbs without noticeable forward movement is an artificial trained figure called "Spanish Walk" [20]. There are likely certain aspects of the forelimb movement which are necessary to the elongation of the stride, and which therefore should be the best indicators for quantifying ground coverage. Another kinematic parameter associated with SL is the over-tracking distance (OTD), the length between the fore hoof imprint and the following imprint of the ipsilateral hind hoof in the direction of travel. OTD explained most of the variance in SL in a sample of ridden Warmblood horses at walk [21]. At walk, breeding experts traditionally estimate OTD by observing the hoof tracks in the arena, due to the gait's specific footfall sequence (left hind, left fore, right hind, right fore) [21]. OTD is also perceived as the gold standard for measuring hind limb engagement: the further the hind limb can move ("engage") under the horse's trunk, the further the hind hoof imprint will lay in front of the previous ipsilateral fore hoof imprint.

In a previous study, seven official FM breeding experts, evaluating all foals and 3-year old horses presented in hand during the annual breeding competitions, scored 24 stallions presented in hand based on a scoring sheet with a total of six traits at walk and eight at trot, among them ground coverage and over-tracking [2]. The scale ranged from one ("undesirable") to nine ("ideal"). The study revealed that the inter-rater reliabilities (quantified with an intra-class correlation coefficient, ICC) for ground coverage and over-tracking were poor (ICC < 0.50), with expert-specific differences in the scale anchoring, and potentially differences in trait interpretation [2]. The aim of the current study was to improve trait definition of ground coverage and over-tracking at walk.

Firstly, it was important to determine which aspects of the forelimb, but also hind limb and upper body kinematics, were the most strongly associated with SL and OTD, and therefore best qualified to quantify ground coverage and over-tracking, analogous to differences in hip hikes or head nods used to quantify lameness [22]. The same 24 FM stallions than in our previous study [2] were equipped with a comprehensive set of markers, and walked and trotted on an instrumented treadmill at different speeds to account for the speed dependence of kinematic parameters. We then determined a subset of kinematic parameters associated with SL and OTD that are measureable not only with OMC, but also IMUs, to make practical recommendations for field applications, where OMC measurements may not always be practicable. Finally, we compared all the kinematic parameters with previously published subjective scores attributed by six FM breeding experts to the 24 measured stallions during an assessment of the video recordings of the treadmill measurements [23], to interpret their scores in relation to objective kinematic measurements. We hypothesise that not all informative parameters are measurable with IMUs, and that the expert scores will show individual tendencies of correlation to specific parameters.

2. Material and Methods

2.1. Animals

In this study, 20 FM stallions from the Swiss National Stud Farm (SNSF) and four old-type FM stallions from private owners were included (mean ± SD; age = 8.8 ± 4.1 years, height at withers = 1.57 ± 0.03 m, and weight = 526.3 ± 32.7 kg). Old-type FM show no Warmblood or Arabian introgression in their pedigree since 1950, and are of a heavier type than the average FM. With both FM and old-type FM, the 24 FM stallions represented the morphological variation in the available FM stallion population at the time of the study [2]. The experimental protocol was approved by the Animal Health and Welfare Commission of the Canton of Vaud (permission number VD 3164). All stallions passed a clinical lameness exam, including a trot up and flexion test performed by the same veterinarian 1 week prior to the beginning of the kinematic measurements. Stallions were habituated to the treadmill with eight to 14 training sessions over 6 weeks before the incremental speed test.

2.2. Data Collection

Stallions were tested at the equine performance laboratory of the veterinary clinic of the University of Zurich on a high-speed instrumented treadmill (Mustang 2200, Ansorix Systems AG, Switzerland) able to measure vertical ground reaction force and hoof position during the stance phase of all four limbs [24]. The positions of the multiple skin mounted spherical reflective makers (SRM) were registered with 10 infra-red OMC cameras (Oqus 7+, Qualisys AB, Sweden). For detailed marker positions see Figure S1. Both systems were synchronized in time, and the stallions were recorded on video during the incremental speed test. The walking speed ranged from 1.4 to 2.0 m/s at 0.1 m/s increments. The common speeds for all stallions were 1.7 and 1.8 m/s. The stallions were measured during 20 seconds for each speed increment. Strides for which measurement values exceeded the standard deviation by a factor three, were excluded. The individual number of strides per speed increment at the walk are summarized in Table S1.

2.3. Data Processing

Spatial parameters measured by the instrumented treadmill (TiF) were calculated with the treadmill software (HP2, University of Zurich, Switzerland) [24]. Marker tracking was done with the Qualisys motion capture software QTM (version 2.9, Qualisys AB, Sweden). Raw 3D coordinates of each SRM were exported into Matlab (version R2020a) and further processed with custom-written scripts to extract specific parameters. Marker interference was filtered out for specific stallions or trials by marking the parameters as missing. Stride segmentation was performed using the hoof-on moments of the left forelimb as previously described [24]. Kinematic analysis was limited to markers of the midline and the left body side [25], as horses were only video-recorded from the left hand side, except for parameters comparing angular differences between contralateral limbs.

2.4. Parameter Selection

The list of putative kinematic parameters for quantifying ground coverage and over-tracking is presented in Table 1. For the trait ground coverage, kinematic parameters related to the forelimbs, hind limbs and pelvis were associated to SL. For the trait over-tracking, only kinematic parameters related to the hind limbs and the pelvis were associated to OTD. For all parameters measured with OMC, we determined whether they could be measured with IMU sensors placed on the head, withers, pelvis and limbs.

2.5. Expert Scoring

The entire incremental speed test (containing walk and trot, not presented here) was filmed from hind, left hand side and front view cameras (HDR-CX760, Sony, Japan). We prepared shorter video sequences of only one common speed (standardised condition for all stallions) and the peak speed to account for individual performance of the stallions and the speed dependence of SL and OTD [27,28]. This allowed us to limit the overall video assessments to 2 hours. The standardised speed for walk was set at 1.7 m/s, at one increment below the peak speed for the slowest stallions. Individual peak speed for walk was determined as the last increment at which the stallions moved regularly with a clear four-beat. The video sequences also contain trot (not discussed here), and were prepared as follows: 20 second clips for the two gaits and speeds were prepared for each stallion from the hind, left hand side and front view cameras (first the two walk sessions followed by the two trot sessions, in order of speed: walk at 1.7 m/s, peak walk, trot at 4.5 m/s, peak trot). The videos are publically available [29]. However, only the walk is considered here.

On two separate occasions, six out of the nine official FM breeding experts, designated by a letter from A to F, appraised the walk and trot of the 24 FM from the prepared video clips [23,29]. The scoring sheet, designed for the appraisal of FM horses in hand [2], contained six traits at walk and eight at trot. The scale ranged from one (“undesirable”) to nine (“ideal”). In this study, only the experts’ scores awarded for ground coverage and over-tracking at walk were considered.

In the first video visualisation round, the experts watched the videos in pairs (A-B, C-D, and E-F). Video clips were assembled in a different order for each pair. In the second video visualisation round, the experts were regrouped based on their native language (French or Swiss-German) in two trios (A-B-C, D-E-F), reappraising the videos in an order specific to each group. Each expert watched the same videos twice but in a different order. Based on the estimated inter-rater reliability of the experts’ scores, the scores from the second visualisation round, with higher intraclass correlation coefficients (ICC) for ground coverage and over-tracking were used for the comparisons with the kinematic parameters [23].

2.6. Statistical Associations

Mean, standard deviations and correlations of the kinematic parameters to SL and/or OT were calculated at each speed increment. Linear mixed effect models (LMEs) were computed based on the formula $Y = X_i\beta + Z_i b_i + \varepsilon_i$, where Y is the outcome variable, X_i is the model matrix for the fixed effects, β is the vector of fixed effects, Z_i is the model matrix for the random effects, b_i is the vector of random effects on each individual i and ε_i is the error term. We computed two models each for ground coverage and over-tracking using the R package *lmerTest* [30].

For ground coverage, we computed a full model, with SL as the predictor variable (Y), all the kinematic variables from Table 1 as fixed effects (β) and the stallion and speed as random effects (b_i). We also computed a reduced model, with SL as the predictor variable (Y), the kinematic variables measureable with IMUs ($Prot_{max}$_MC, Ret_{max}_MC, $Prot_{max}$_MT, Ret_{max}_MT, CLProRet_MC, CLProRet_MT, Z_{ROM}_pelv, $Pitch_{max}$_pelv, $Pitch_{ROM}$_pelv, Yaw_{max}_pelv, Yaw_{ROM}_pelv, $Roll_{max}$_pelv, $Roll_{ROM}$_pelv) as fixed effects (β), and the stallion and speed as random effects (b_i).

For over-tracking, we computed a full model with OT as predictor variable, hind limb and pelvis parameters as fixed effects (β) and stallion and speed as random effects (b_i). The reduced model consisted of OT as the predictor variable (Y), metatarsal and pelvis parameters as fixed effects (β) and stallion and speed as random effects (b_i).

For each of the four models, the collinearity of the parameters was checked using the package *performance*, and parameters with a variance inflation factor (VIF) > 10 were removed from the model [31]. We also estimated the effect size of each parameter using partial Eta-squared (η^2), quantifying the proportion of variance in the model explained by the parameter, using the R package *effectsize* on the fixed effect parameters [32]. Effect sizes (η^2) > 0.04 mean an explained variance of 4%, and can be considered as medium effects, η^2 > 0.14 as large [33]. The scores from each expert were correlated to all kinematic parameters at each of the two speeds (standard and peak) separately using Pearson’s correlation coefficients. Correlations of expert’s scores to a parameter equal to or above 0.41 were considered significant in absolute values for 24 horses, following Krehbiel’s rule of thumb [34].

3. Results

Stride length (SL) and over-tracking distance (OTD), the two outcome parameters, increased linearly with speed. Descriptive statistics and correlation between SL and putative parameters affecting ground coverage at each walking speed for the forelimbs are summarized in Table S2. Descriptive statistics for the hind limb parameters at each speed increment are presented in Table S3, while correlations between hind limb parameters and SL as well as OTD are detailed in Table S4. The inter-correlations of all parameters measured at the standard speed of 1.7 m/s were visualised in a cross-correlation matrix (Fig. 1).

3.1. Association Between Fore and Hind Limb Kinematic Parameters and Stride Length

The parameters distance of the hoof marker to the tuber coxae marker (Hoof-TC), the contralateral pro-retraction angle of the forelimbs (CLProRet_front) and the maximal hind limb retraction angle (Ret_{max}_hind) were excluded from the LME due to high VIF. The coefficient of determination of the LME was very high

Table 1
Definition of parameters and their abbreviations putatively associated to ground coverage (GC) or over-tracking (OT) measured with the instrumented treadmill (TiF) and optical motion capture (OMC), with a mention of which parameters can be measured with inertial measurement units (IMU). The axes are defined as in Clayton and Hobbs [26]. All values are the mean over all available strides unless otherwise stated.

Parameter	Definition	Units	Trait	Measurable With IMUs
Spatial parameters				
SL	Stride length; derived from the stride duration based on the hoof-on moments for the left front hoof and the speed of the treadmill	[m]	GC	yes
OTD	Over-tracking distance of the left hind hoof in relation to the left front hoof	[m]	GC/OT	no
Hoof-TC	Horizontal distance of the left hind hoof relative to the vertical from the ipsilateral tuber coxae during hind limb protraction	[m]	GC/OT	no
Forelimb parameters				
Prot$_{max}$_front	Maximum protraction angle of the left forelimb (marker tuber spina scapula to fetlock in relation to the vertical)	[deg]	GC	no
Prot$_{max}$_MC	Maximal metacarpus protraction angle of the left forelimb (cluster rotation around the transverse axis)	[deg]	GC	yes
Ret$_{max}$_front	Maximal retraction angle of the left forelimb (marker tuber spina scapula to fetlock in relation to the vertical, negative value)	[deg]	GC	no
Ret$_{max}$_MC	Maximal metacarpus retraction angle of the left forelimb (cluster rotation around the transverse axis, negative value)	[deg]	GC	yes
Prot$_{height}$_front	Maximal limb protraction height of the left forelimb, normalised for withers height. Vertical position of the hoof marker relative to the ground	[m]	GC	no
Prot$_{height}$@Prot$_{max}$_front	Maximum limb protraction height at maximal protraction angle of the left forelimb, normalised for withers height	[m]	GC	no
A$_{fetlock}$_front	Maximum fetlock hyperextension angle of the left forelimb during midstance	[deg]	GC	no
CLProRet_front	Absolute maximal difference in the protraction-retraction angles of contralateral forelimbs, combined FL-FR, FR-FL	[deg]	GC	no
CLProRet_MC	Absolute maximal difference in the protraction-retraction angles of contralateral metacarpi, combined FL-FR, FR-FL	[deg]	GC	yes
Yaw$_{ROM}$_forehand	Range of forehand yaw. Range of rotation of the L/R tuber spina scapulae vector around the vertical axis, corrected for the longitudinal orientation of the trunk (virtual vector from the centre of the tuber spina scapulae to S6)	[deg]	GC	no
Hind limb parameters				
Prot$_{max}$_hind	Maximum protraction angle of the left hind limb (hip to fetlock marker in relation to the vertical)	[deg]	GC/OT	no
Prot$_{max}$_MT	Maximal metatarsus protraction angle for left hind limb (cluster rotation around the transverse axis)	[deg]	GC/OT	yes
Ret$_{max}$_hind	Maximal retraction angle of the left hind limb (hip to fetlock marker in relation to the vertical, negative value)	[deg]	GC/OT	no
Ret$_{max}$_MT	Maximal metatarsus retraction angle for left hind limb (cluster rotation around the transverse axis, negative value)	[deg]	GC/OT	yes
A$_{fetlock}$_hind	Maximum fetlock hyperextension angle from the left hind limb during midstance (angles between hoof – fetlock and MC cluster markers)	[deg]	GC/OT	no
CLProRet_hind	Absolute maximal difference in the protraction-retraction angles of contralateral hind limbs, combined HL-HR, HR-HL	[deg]	GC/OT	no
CLProRet_MT	Absolute maximal difference in the protraction-retraction angles of contralateral metatarsi, combined HL-HR, HR-HL	[deg]	GC/OT	yes
Pelvis parameter				
Z$_{ROM}$_pelv	Vertical range of movement of the S1 marker (from the pelvis marker cluster)	[m]	GC/OT	yes
Pitch$_{max}$_pelv	Maximum pelvis pitch calculated as the rotation of the S1-S6 vector (from the pelvis marker cluster) around the transverse axis of the horse	[deg]	GC/OT	yes
Pitch$_{ROM}$_pelv	Range of pelvis pitch calculated as the rotation of the S1-S6 vector around the transverse axis of the horse	[deg]	GC/OT	yes
Yaw$_{max}$_pelv	Maximum pelvis yaw calculated as the rotation of the S1-S6 vector around the vertical axis	[deg]	GC/OT	yes
Yaw$_{ROM}$_pelv	Range of pelvis yaw calculated as the rotation of the S1-S6 vector around the vertical axis	[deg]	GC/OT	yes
Roll$_{max}$_pelv	Maximum pelvis roll calculated as the rotation of the L and R sacrum vector around the longitudinal axis	[deg]	GC/OT	yes
Roll$_{ROM}$_pelv	Range of pelvis roll calculated between the L and R sacrum vector around the longitudinal axis	[deg]	GC/OT	yes

(R^2=0.98), with 11 out of 22 tested kinematic parameters significantly associated with stride length (SL; Table 2). Five parameters had medium effect sizes ($\eta^2 > 0.04$): the maximal forelimb retraction angle (Ret$_{max}$_front), the range of motion of the yaw of the forehand (Yaw$_{ROM}$_forehand), the maximal forelimb protraction angle (Prot$_{max}$_front) and the maximal fetlock hyperextension angle (A$_{fetlock}$_front) explained 11%, 6%, 5%, and 5% of the variance in SL, respectively.

3.2. Stride Length Model Based on Movement Parameters Measurable With IMUs

A reduced LME was computed with the OMC parameters that can also be measured with IMUs. The collinearity between parameters was low and none had to be excluded from the LME beforehand. The coefficient of determination of the IMU-specific LME was as high as the first model ($R^2 = 0.98$), with 6 out of 13 tested kinematic parameters significantly associated with SL (Table 3). Three parameters had medium effect sizes ($\eta^2 > 0.04$): the range of motion of pelvis pitch (Pitch$_{ROM}$_pelv), the maximal metacarpal protraction angle (Prot$_{max}$_MC) and the maximal pelvis pitch (Pitch$_{max}$_pelv), explaining 10%, 8% and 6% of the variance in SL, respectively, and when considering only parameters measurable with IMUs.

3.3. Correlations Between Subjective Ground Coverage Scores and Objective Kinematic Measurements

The exact correlations between the subjective ground coverage scores and the objective kinematic measurements at standard (1.7 m/s) and peak individualised speed (1.8–2.0 m/s) are presented in the supplementary Table S5, and summarised in Table 2 and Table 3. Ground coverage scores from experts A, B, C and F were positively correlated to SL at standard speed but only the scores

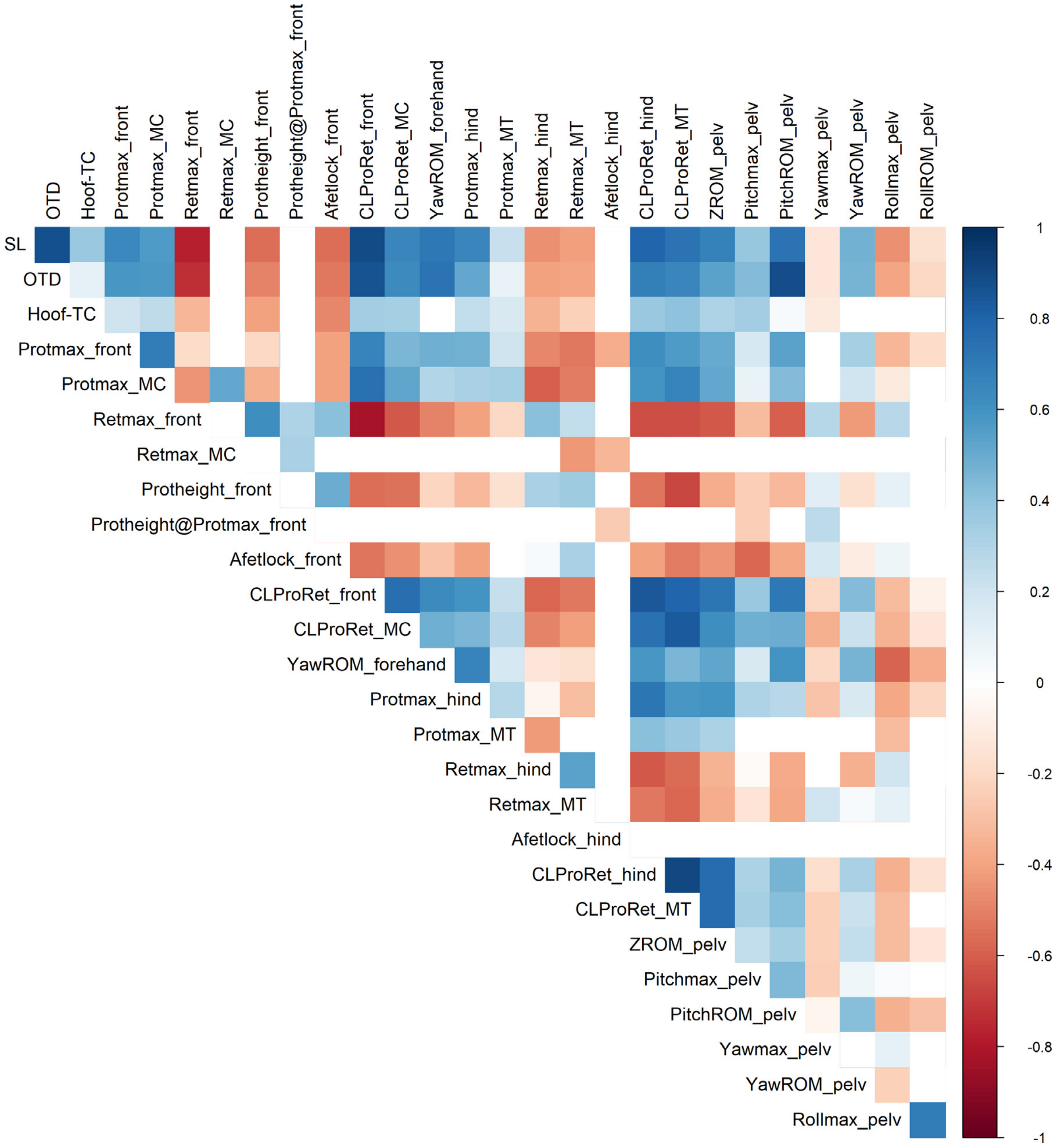

Fig. 1. Cross-correlation matrix of all movement parameters, measured at 1.7 m/s. SL, stride length; OTD, over-tracking distance; Hoof-TC, distance between hoof and tuber coxae during hind limb protraction; Protmax_front, maximal forelimb protraction; Protmax_MC, maximal metacarpal protraction; Retmax_front, maximal forelimb retraction; Retmax_MC, maximal metacarpal retraction; Protheight_front, maximal forelimb protraction height; Protheight@Protmax_front, maximal forelimb protraction height at maximal protraction; Afetlock_front, maximum forelimb fetlock hyperextension; CLProRet_front, contralateral pro-retraction angles forelimbs; CLProRet_MC, contralateral metacarpal pro-retraction angles; YawROM_forehand, range of yaw of the forehand; Protmax_hind, maximal hind limb protraction; Protmax_MT, maximal metatarsal protraction; Retmax_hind, maximal hind limb retraction; Retmax_MT, maximal metatarsal retraction; CLProRet_hind, contralateral pro-retraction angles hind limbs; CLProRet_MT, contralateral metatarsal pro-retraction angles; ZROM_pelv, vertical range of motion pelvis; Pitchmax_pelv, maximal pelvis pitch; PitchROM_pelv, range of motion pelvis pitch; Yawmax_pelv, maximal pelvis yaw; YawROM_pelv, range of motion pelvis yaw; Rollmax_pelv, maximal roll pelvis; RollROM_pelv, range of motion pelvis yaw.

from experts A and D were correlated with SL at peak speed. The scores from Expert C had the highest correlation with any kinematic parameter ($r = 0.85$ with Prot$_{max}$_front at standard speed). In total, 19 out of 22 parameters included in the full OMC model were correlated $r \geq |0.41|$ with the ground coverage score of at least one expert for either standard (16 parameters) or peak speed (15 parameters) (Table 2). The scores from experts A, B and C correlated to most of the same parameters. The scores from experts D, E, and F did not show particular patterns of association with those from other experts.

The kinematic parameter with the highest effect size in the OMC model, Ret$_{max}$_front, was negatively correlated to the scores from experts A, B and C, but only at standard speed of 1.7 m/s. The scores from all experts were negatively correlated to the A$_{fetlock}$_front at standard speed (expert F), peak speed (experts D and E) or both (experts A, B, and C), that is ground coverage scores were higher when Ret$_{max}$_front and A$_{fetlock}$_front were larger (as the retraction and extension angles were on a negative axis). 12 out of 16 parameters correlating with ground coverage scores at standard speed, and 13 out of 15 parameters at peak speed, explained less than 5% of the SL variance.

Table 2

Associations between kinematic parameters and stride length, with speed (1.4 m/s to 2.0 m/s) and horse ($n = 24$) as random factors, with the effect sizes η^2 and their 95% confidence intervals (CI). The individual scores for ground coverage by the experts (defined by a letter from A to F) correlated above the significance threshold of $r \geq |0.41|$ to a parameter are reported in the last two columns, with a superscript for the direction of correlation. Parameters which were significantly correlated ($r \geq |0.41|$) to three or more individual expert scores are in bold. Parameter abbreviations are defined in Table 1.

Parameter	F-value	*P*-value	Effect Size η^2	CI	Expert $r \geq |0.41|$ at 1.7 m/s	Expert $r \geq |0.41|$ at Peak Speed
Ret$_{max}$_front	**149.16**	**<.0001**	**0.11**	**[0.08, 0.15]**	**A^-, B^-, C^-**	
Yaw$_{ROM}$_forehand	69.79	<.0001	0.06	[0.04, 0.09]	C^+	
Prot$_{max}$_front	**57.96**	**<.0001**	**0.05**	**[0.03, 0.07]**	**A^+, B^+, C^+**	**A^+, B^+, C^+, E^+**
A$_{fetlock}$_front	**13.89**	**<.0001**	**0.05**	**[0.01, 0.10]**	**A^-, B^-, C^-, F^-**	**A^-, B^-, C^-, D^-, E^-**
Pitch$_{ROM}$_pelv	**41.70**	**<.0001**	**0.04**	**[0.02, 0.07]**	**A^+, B^+, C^+, F^+**	**B^+**
OTD	**30.90**	**<.0001**	**0.03**	**[0.01, 0.05]**	**A^+, B^+, C^+, F^+**	**A^+**
Prot$_{max}$_hind	36.02	<.0001	0.03	[0.01, 0.05]	C^+	C^+
Ret$_{max}$_MC	22.36	<.0001	0.02	[0.01, 0.04]		B^+
Prot$_{max}$_MT	19.19	<.0001	0.02	[0.00, 0.03]		F^+
Ret$_{max}$_MT	**13.36**	**.0003**	**0.01**	**[0.00, 0.03]**	**A^-, B^-, C^-**	**A^-, B^-, C^-, E^-**
Prot$_{max}$_MC	**4.68**	**.0307**	**3.96E-03**	**[0.00, 0.01]**	**A^+, B^+, C^+, F^+**	**A^+, B^+, C^+, E^+**
Roll$_{ROM}$_pelv	2.15	.1428	1.83E-03	[0.00, 0.01]		
Yaw$_{ROM}$_pelv	1.29	.2567	1.09E-03	[0.00, 0.01]	B^+	
Roll$_{max}$_pelv	0.69	.4055	5.92E-04	[0.00, 0.01]		
Yaw$_{max}$_pelv	0.43	.5109	3.70E-04	[0.00, 0.01]		
Prot$_{height}$_front	0.41	.5216	3.50E-04	[0.00, 0.01]	F^-	
CLProRet_MT	**0.39**	**.5302**	**3.49E-04**	**[0.00, 0.01]**		**A^+, C^+, D^+, E^+, F^+**
Z$_{ROM}$_pelv	**0.23**	**.6335**	**1.94E-04**	**[0.00, 0.00]**	**A^+, B^+, C^+**	**A^+, D^+, E^+**
Pitch$_{max}$_pelv	0.05	.8229	1.91E-04	[0.00, 0.01]	C^+, F^+	D^+
CLProRet_MC	**0.14**	**.7065**	**1.21E-04**	**[0.00, 0.00]**	**B^+, C^+, F^+**	**D^+, E^+, F^+**
A$_{fetlock}$_hind	**0.03**	**.8698**	**7.12E-05**	**[0.00, 0.01]**		**B^+, C^+, D^+, F^+**
Prot$_{height}$@Prot$_{max}$_front	**0.03**	**.8736**	**2.16E-05**	**[0.00, 0.00]**	**E^+**	**A^+, B^+, C^+, E^+**

Abbreviations: A$_{fetlock}$_front, maximum forelimb fetlock hyperextension; A$_{fetlock}$_hind, Maximum fetlock hyperextension; CLProRet_MC, contralateral metacarpal pro-retraction angles; CLProRet_MT, contralateral metatarsal pro-retraction angles; OTD, over-tracking distance; Pitch$_{max}$_pelv, maximal pelvis pitch; Pitch$_{ROM}$_pelv, range of motion pelvis pitch; Prot$_{height}$_front, maximal forelimb protraction height; Prot$_{height}$@Prot$_{max}$_front, maximal forelimb protraction height at maximal protraction; Prot$_{max}$_front, maximal forelimb protraction; Prot$_{max}$_hind, maximal hind limb protraction; Prot$_{max}$_MC, maximal metacarpal protraction; Prot$_{max}$_MT, maximal metatarsal protraction; Ret$_{max}$_front, maximal forelimb retraction; Ret$_{max}$_MC, maximal metacarpal retraction; Ret$_{max}$_MT, maximal metatarsal retraction; Roll$_{max}$_pelv, maximal roll pelvis; Roll$_{ROM}$_pelv, range of motion pelvis yaw; Yaw$_{max}$_pelv, maximal pelvis yaw; Yaw$_{ROM}$_forehand, range of yaw of the forehand; Yaw$_{ROM}$_pelv, range of motion pelvis yaw; Z$_{ROM}$_pelv, vertical range of motion pelvis.

Table 3

Associations between kinematic parameters measurable with IMUs and stride length, with speed (1.4 m/s to 2.0 m/s) and horse ($n = 24$) as random factors, with the effect sizes η^2 and their 95% confidence intervals (CI). The individual scores for ground coverage by the experts (defined by a letter from A to F) correlated above the significance threshold of $r \geq |0.41|$ to a parameter are reported in the last two columns, with a superscript for the direction of correlation. Parameters which were significantly correlated ($r \geq |0.41|$) to three or more individual expert scores are in bold. Parameter abbreviations are defined in Table 1.

Parameter	F-value	*P*-value	Effect Size η^2	CI	Expert $r > |0.41|$ at 1.7 m/s	Expert $r > |0.41|$ at Peak Speed
Pitch$_{ROM}$_pelv	**119.89**	**<.0001**	**0.10**	**[0.07, 0.13]**	**A^+, B^+, C^+, F^+**	**B^+**
Prot$_{max}$_MC	**107.06**	**<.0001**	**0.08**	**[0.05, 0.11]**	**A^+, B^+, C^+, F^+**	**A^+, B^+, C^+, E^+**
Pitch$_{max}$_pelv	32.49	<.0001	0.06	[0.03, 0.10]	C^+, F^+	D^+
Yaw$_{ROM}$_pelv	27.76	<.0001	0.02	[0.01, 0.04]	B^+	
CLProRet_MC	**18.85**	**<.0001**	**0.01**	**[0.00, 0.03]**	**B^+, C^+, F^+**	**D^+, E^+, F^+**
CLProRet_MT	**17.03**	**<.0001**	**0.01**	**[0.00, 0.03]**		**A^+, C^+, D^+, E^+, F^+**
Ret$_{max}$_MC	4.58	.0326	3.61E-03	[0.00, 0.01]		B^+
Prot$_{max}$_MT	4.47	.0347	3.54E-03	[0.00, 0.01]		F^+
Roll$_{max}$_pelv	1.78	.1829	1.40E-03	[0.00, 0.01]		
Yaw$_{max}$_pelv	1.45	.2289	1.16E-03	[0.00, 0.01]		
Roll$_{ROM}$_pelv	1.30	.2536	1.04E-03	[0.00, 0.01]		
Z$_{ROM}$_pelv	**0.52**	**.4696**	**4.16E-04**	**[0.00, 0.01]**	**A^+, B^+, C^+**	**A^+, D^+, E^+**
Ret$_{max}$_MT	**0.09**	**.7621**	**7.25E-05**	**[0.00, 0.00]**	**A^-, B^-, C^-**	**A^-, B^-, C^-, E^-**

Abbreviations: CLProRet_MC, contralateral metacarpal pro-retraction angles; CLProRet_MT, contralateral metatarsal pro-retraction angles; IMUs, inertial measurement units; Pitch$_{max}$_pelv, maximal pelvis pitch; Pitch$_{ROM}$_pelv, range of motion pelvis pitch; Prot$_{max}$_MC, maximal metacarpal protraction; Prot$_{max}$_MT, maximal metatarsal protraction; Ret$_{max}$_MC, maximal metacarpal retraction; Ret$_{max}$_MT, maximal metatarsal retraction; Roll$_{max}$_pelv, maximal roll pelvis; Roll$_{ROM}$_pelv, range of motion pelvis yaw; Yaw$_{max}$_pelv, maximal pelvis yaw; Yaw$_{ROM}$_pelv, range of motion pelvis yaw; Z$_{ROM}$_pelv, vertical range of motion pelvis.

When only considering the parameters measured with IMUs, a larger Pitch$_{ROM}$_pelv and Pitch$_{max}$_pelv, Prot$_{max}$_MC and contralateral metacarpal pro-retraction angle (CLProRet_MC) were positively and significantly correlated with ground coverage scores by a majority of experts at standard speed. At peak speed, the contralateral metatarsal pro-retraction angle (CLProRet_MT) was positively correlated to ground coverage scores given by five out of the six experts, despite its lower effect size. Two parameters, the vertical range of motion of the pelvis (Z$_{ROM}$_pelv) and the maximal metatarsal retraction angle (Ret$_{max}$_MT) were correlated to scores by at least three experts at either speed, but these parameters were not significant in the IMU model.

3.4. Association Between Hind Limb Parameters and Over-Tracking Distance

The parameters Hoof-TC and Ret$_{max}$_hind were excluded from the model due to high VIF. The coefficient of determination of the final LME was relatively high ($R^2 = 0.93$), but lower than for SL, and 8 out of 15 tested kinematic parameters significantly associated with OTD (Table 4). One parameter, Pitch$_{ROM}$_pelv, had a high ($\eta^2 > 0.14$) effect size, explaining 24% of the variance in OTD in this model, while two additional parameters had medium effect sizes ($\eta^2 > 0.04$): A$_{fetlock}$_front and Prot$_{max}$_hind, explaining 11% and 7% of the variance in the OTD model, respectively.

Table 4
Associations between kinematic parameters and over-tracking distance, with speed (1.4 m/s to 2.0 m/s) and horse (n = 24) as random factors, with the effect sizes η^2 and their 95% confidence intervals (CI). The individual scores for over-tracking by the experts (defined by a letter from A to F) correlated above the significance threshold of $r \geq |0.41|$ to a parameter are reported in the last two columns, with a superscript for the direction of correlation. Parameters which were significantly correlated ($r \geq |0.41|$) to three or more individual expert scores are in bold. Parameter abbreviations are defined in Table 1.

Parameter	F-value	*P*-value	Effect Size η^2	CI	Expert r > |0.41| at 1.7 m/s	Expert r > |0.41| at Peak Speed
Pitch$_{ROM}$_pelv	168.40	<.0001	0.24	[0.18, 0.30]	A$^+$, C$^+$	
A$_{fetlock}$_hind	9.83	.0024	0.11	[0.01, 0.25]	D$^-$	A$^-$, D$^-$
Prot$_{max}$_hind	82.49	<.0001	0.07	[0.05, 0.11]	A$^+$, C$^+$	C$^+$
Prot$_{max}$_MT	28.54	<.0001	0.03	[0.01, 0.06]		
Ret$_{max}$_MT	**23.45**	**<.0001**	**0.03**	**[0.01, 0.05]**	**A$^-$, B$^-$, C$^-$**	**A$^-$, E$^-$**
CLProRet_hind	**20.65**	**<.0001**	**0.03**	**[0.01, 0.07]**	**A$^+$, B$^+$, C$^+$**	**A$^+$, C$^+$, D$^+$**
Pitch$_{max}$_pelv	**1.99**	**.1625**	**0.03**	**[0.00, 0.14]**	**A$^+$, C$^+$, D$^+$**	
Roll$_{max}$_pelv	23.84	<.0001	0.02	[0.01, 0.04]		
Yaw$_{ROM}$_pelv	**16.62**	**<.0001**	**0.01**	**[0.00, 0.03]**	**B$^+$, C$^+$, D$^+$**	**D$^+$**
Yaw$_{max}$_pelv	3.99	.0460	3.51E-03	[0.00, 0.01]		
CLProRet_MT	**0.22**	**.6366**	**4.49E-04**	**[0.00, 0.01]**	**A$^+$, C$^+$, F$^+$**	**A$^+$, C$^+$, D$^+$, E$^+$**
Z$_{ROM}$_pelv	0.38	.5363	3.41E-04	[0.00, 0.01]	A$^+$, C$^+$	D$^+$
Roll$_{ROM}$_pelv	<0.01	.9811	5.00E-07	[0.00, 0.00]	B$^-$	

Abbreviations: CLProRet_MC, contralateral metacarpal pro-retraction angles; CLProRet_MT, contralateral metatarsal pro-retraction angles; Pitch$_{max}$_pelv, maximal pelvis pitch; Pitch$_{ROM}$_pelv, range of motion pelvis pitch; Prot$_{max}$_MC, maximal metacarpal protraction; Prot$_{max}$_MT, maximal metatarsal protraction; Ret$_{max}$_MC, maximal metacarpal retraction; Ret$_{max}$_MT, maximal metatarsal retraction; Roll$_{max}$_pelv, maximal roll pelvis; Roll$_{ROM}$_pelv, range of motion pelvis yaw; Yaw$_{max}$_pelv, maximal pelvis yaw; Yaw$_{ROM}$_pelv, range of motion pelvis yaw; Z$_{ROM}$_pelv, vertical range of motion pelvis.

Table 5
Associations between kinematic parameters measurable with IMUs in the field and over-tracking distance, with speed (1.4 m/s to 2.0 m/s) and horse (n = 24) as random factors, with the effect sizes η^2 and their 95% confidence intervals (CI). The individual scores for over-tracking by the experts (defined by a letter from A to F) correlated above the significance threshold of $r \geq |0.41|$ to a parameter are reported in the last two columns, with a superscript for the direction of correlation. Parameters which were significantly correlated ($r \geq |0.41|$) to three or more individual expert scores are in bold. Parameter abbreviations are defined in Table 1.

Parameter	F-value	*P*-value	Effect Size η^2	CI	Expert r > |0.41| at 1.7 m/s	Expert r > |0.41| at Peak Speed
Pitch$_{ROM}$_pelv	246.44	<.0001	0.33	[0.27, 0.39]	A$^+$, C$^+$	
Pitch$_{max}$_pelv	**8.75**	**.0042**	**0.11**	**[0.01, 0.26]**	**A$^+$, C$^+$, D$^+$**	
CLProRet_MT	**106.34**	**<.0001**	**0.10**	**[0.07, 0.13]**	**A$^+$, C$^+$, F$^+$**	**A$^+$, C$^+$, D$^+$, E$^+$**
Ret$_{max}$_MT	**42.98**	**<.0001**	**0.04**	**[0.02, 0.07]**	**A$^-$, B$^-$, C$^-$**	**A$^-$, E$^-$**
Yaw$_{ROM}$_pelv	**44.93**	**<.0001**	**0.03**	**[0.01, 0.05]**	**B$^+$, C$^+$, D$^+$**	**D$^+$**
Roll$_{max}$_pelv	13.87	.0002	0.01	[0.00, 0.03]		
Roll$_{ROM}$_pelv	9.04	.0027	7.149E-03	[0.00, 0.02]	B$^-$	
Yaw$_{max}$_pelv	1.64	.2011	1.29E-03	[0.00, 0.01]		
Prot$_{max}$_MT	1.04	.3075	1.05E-03	[0.00, 0.01]		
Z$_{ROM}$_pelv	1.26	.2618	9.93E-04	[0.00, 0.01]	A$^+$, C$^+$	D$^+$

Abbreviations: CLProRet_MT, contralateral metatarsal pro-retraction angles; IMUs, inertial measurement units; Pitch$_{max}$_pelv, maximal pelvis pitch; Pitch$_{ROM}$_pelv, range of motion pelvis pitch; Prot$_{max}$_MT, maximal metatarsal protraction; Ret$_{max}$_MT, maximal metatarsal retraction; Roll$_{max}$_pelv, maximal roll pelvis; Roll$_{ROM}$_pelv, range of motion pelvis yaw; Yaw$_{max}$_pelv, maximal pelvis yaw; Yaw$_{ROM}$_pelv, range of motion pelvis yaw; Z$_{ROM}$_pelv, vertical range of motion pelvis.

3.5. Over-Tracking Distance Model Based on Hind Limb Parameters Measurable With IMUs

A reduced LME was computed with the hind limb and pelvic parameters that can be measured with IMUs. The collinearity between parameters was low and none had to be excluded from the LME. The coefficient of determination of the LME was $R^2 = 0.93$. Seven out of 10 tested kinematic parameters were significantly associated with OTD (Table 5). Pitch$_{ROM}$_pelv had a high ($\eta^2 > 0.14$) effect size explaining 33% of the variance in OTD in this model, while two additional parameters had medium effect sizes ($\eta^2 > 0.04$): Pitch$_{max}$_pelv and CLProRet_MT, explaining 11% and 10% of the variance in the OTD model, respectively.

3.6. Correlations Between Subjective Over-Tracking Scores and Objective Kinematic Measurements

The over-tracking scores from all experts except expert E were significantly correlated with OTD at either standard speed (experts A, B, C, D, F) or peak individualised speed (experts A and D). The scores from Expert C showed the highest correlation ($r = 0.75$) with the contralateral pro-retraction angle of the hind limbs (CLProRet_hind). The score from expert E was only correlated with the measures at peak speed.

The scores from only two experts were significantly correlated with Pitch$_{ROM}$_pelv, the parameter explaining the highest variance in OTD, exclusively at standard speed. The parameter correlating with the most expert scores was CLProRet_MT, although it was not significantly correlated to OTD in the full model. In total, 10 out of 13 parameters were correlated with the over-tracking score of at least one expert for either standard (10 parameters) or peak speed (7 parameters) (Table 4). The scores from all experts except B were correlated to at least one of the parameters with a medium to high effect size. Half of the parameters correlating with over-tracking scores at standard speed, and 5 out of 7 parameters at peak speed, explained less than 5% of the OTD variance.

4. Discussion

In this study, we confirmed that at walk, ground coverage is related both to stride length (SL), and the movement of the forehand. A larger maximal forelimb retraction angle (Ret$_{max}$_front), a larger maximal fetlock hyperextension (A$_{fetlock}$_front), a larger yaw of the forehand (Yaw$_{ROM}$_forehand), and a larger maximal forelimb protraction angle (Prot$_{max}$_front) were associated with an increased stride length. However, none of these parameters can be measured using IMUs. A longer stride was also associated with a larger maximum and range of motion of pelvis pitch (Pitch$_{max}$_pelv and Pitch$_{ROM}$_pelv), and maximal protraction angle of the metacarpus (Prot$_{max}$_MC). The over-tracking distance was best quantified with the range of pelvis pitch (Pitch$_{ROM}$_pelv).

4.1. Ground Coverage

The kinematic parameters with the highest effect sizes for SL were Ret_{max}_front, Yaw_{ROM}_forehand, $\text{A}_{\text{fetlock}}$_front and Prot_{max}_front, all measureable only with OMC. These parameters indicate two mechanisms determining SL at the walk: the movement of extending the forelimb forward together with a yawing movement of the shoulders (Prot_{max}_front and Yaw_{ROM}_forehand), and pivoting the mass of the trunk over the supporting forelimb (increased Ret_{max}_front), compressing the fetlock (increased $\text{A}_{\text{fetlock}}$_front), acting as a vaulting pole. This is representative of the inverted pendulum mechanics of the walk. The importance of Prot_{max}_front and Yaw_{ROM}_forehand were congruent with our hypothesised textual definition of ground coverage, and scores from experts A, B, and C were all substantially correlated to the maximal protraction angle at both standard and peak speed.

The Ret_{max}_front explained more of the SL variance at the walk in the full model than Prot_{max}_front, which is the opposite result from a kinematic study performed on Dutch Warmblood horses [13], where the forelimb protraction angle was considered more important. This discordance can more likely be attributed to methodological differences, as the FM horses were measured over a range of speeds instead of one standard speed (1.7 m/s), rather than to a breed specific difference. Taken together, the range of forelimb retraction-protraction and the maximum retraction-protraction angles showed high heritabilities (0.61 ± 0.26 and 0.59 ± 0.27, respectively) in 130 Pura Raza Espanol stallions measured on a treadmill at 1.7 m/s [16]. Both studies [13,16] were only performed at one speed on different breeds than the FM. Nevertheless, there is now ample evidence that Prot_{max}_front and Ret_{max}_front are important in horse breeding, and are potentially heritable. However, for the FM, none of the experts' scores were substantially correlated to Ret_{max}_front. In the future, putting more emphasis on also observing this parameter could be helpful to improve scoring validity at the walk.

In a field measurement condition, when the limb protraction and retraction angles have to be approximated using the metacarpal and metatarsal angles, and neither the yaw of the forehand nor the fetlock extension angles can be accurately quantified, SL could be modelled using $\text{Pitch}_{\text{ROM}}$_pelv, $\text{Pitch}_{\text{max}}$_pelv, and Prot_{max}_MC. The coefficient of determination for an IMU-specific model was equal to the full model, indicating that SL can still be accurately predicted with a reduced amount of parameters. This means that in the future, horses could be accurately phenotyped for ground coverage in the field using IMU sensors placed on the head, withers, pelvis and limbs, especially if the measurements can be related to speed, as for example proposed in Darbandi et al [35] using machine learning approaches.

4.2. Over-Tracking

The over-tracking distance explained less of SL variance (3%) than we expected based on a previous study of a smaller subset of ridden Warmblood horses [21]. However, the range of pelvis pitch ($\text{Pitch}_{\text{ROM}}$_pelv) was also the most important parameter quantifying OTD for the full model. While the ground coverage and over-tracking scores from experts A, B and C were correlated with $\text{Pitch}_{\text{ROM}}$_pelv at standard speed, the over-tracking scores did not correlate with this parameter at peak speed. Furthermore, scores from experts D, E, and F were never significantly correlated to this parameter. Similarly to the maximal forelimb retraction angle being a good indicator for ground coverage, the range of pelvis pitch seems an excellent indicator for over-tracking, which was not yet considered by all experts (according to the correlations between over-tracking and $\text{Pitch}_{\text{ROM}}$_pelv). In the full model, $\text{A}_{\text{fetlock}}$_hind had the second largest effect on OTD, but not on SL. The importance of $\text{A}_{\text{fetlock}}$_hind in this model is somewhat unclear in this study, considering the low correlations of this parameter to the others. However, both $\text{A}_{\text{fetlock}}$_front and $\text{A}_{\text{fetlock}}$_hind parameters were positively correlated to the gait quality score in Dutch Warmblood horses in absolute values [13], albeit better associated to the trait "suppleness" than to SL. Regrettably, despite its potential for quantifying even more gait quality traits, fetlock hyperextension cannot currently be derived reliably from IMUs. In the OMC-specific LME, Prot_{max}_hind had a medium effect size, which is consistent with the textual definition of over-tracking.

For the IMU-specific model, $\text{Pitch}_{\text{ROM}}$_pelv, $\text{Pitch}_{\text{max}}$_pelv, CL-ProRet_MT and Ret_{max}_MT had the highest effect sizes. The overwhelming effect of $\text{Pitch}_{\text{ROM}}$_pelv on OTD explains that both models had equal coefficients of determination ($R^2 = 0.93$). Considering the IMU-specific model for SL as well, $\text{Pitch}_{\text{ROM}}$_pelv is important at the walk to increase both OTD and SL. Conveniently, $\text{Pitch}_{\text{ROM}}$_pelv is easily measurable with an IMU on the sacrum, and could be an interesting, objective phenotype for future genetic studies. However, the coefficients of determination for OTD were lower than for SL ($R^2 = 0.98$), suggesting that other parameters relating to the forelimb which we excluded from the models, may also have an influence on OTD.

4.3. Scoring Tendencies of the Experts

As expected from the poor inter-rater reliability estimates between the experts (ground coverage: ICC = 0.34, over-tracking: ICC = 0.21) [23], they showed different tendencies in their scorings. The scores from experts A, B and C were highly correlated [23] and accordingly these three experts had similar tendencies in relation to the kinematic parameters. The scorings of experts A, B, C and also F were correlated significantly ($>|0.41|$) to more parameters at standard speed than peak speed. This might suggest that they scored mostly based on the first part of the video clips (the standard speed). In contrast, expert E had consistently higher correlations to more parameters at peak speed in comparison to standard speed. Expert D was more correlated to hind limb parameters at the walk, and also more at peak speed than standard speed. The effect of the language of the experts and the order of the videos on the scoring is unclear: experts A, B and C (German-speaking, first group) showed similar scoring tendencies, while experts D, E, and F (French-speaking, second group) did not.

While not directly comparable, the inter-rater reliability of these six experts was lower overall than the inter-rater reliability of seven experts (including the six experts from this study) scoring the same horses under classical conditions in a triangle [2]. Furthermore, the inter-rater reliabilities increased from the first to the second scoring, suggesting a form of habituation to scoring videos of horses on the treadmill [23]. However, the highest intra-rater reliabilities at walk for ground coverage (ICC = 0.49, expert E) and over-tracking (ICC = 0.43, expert E) were both still poor [23]. Furthermore, the expert with the highest intra-rater reliabilities (expert E) was not the one with the highest correlations to the objective kinematic measurements (expert C). All these factors suggest that scoring horses walking on the treadmill was difficult for the breeding experts. Based on their verbal feedback, seeing the horses on the treadmill, the lack of background differentiation in movement and the length of the visualisation were major factors in their discomfort scoring the videos.

In this study, it could only be determined that the expert scores correlated more or less to specific parameters depending on the expert, but it is not possible to ascertain whether these were the parameters the experts actually observed. In a future study, the horses could be presented in hand, in a manner familiar to the experts, equipped with IMUs. The experts could be equipped with goggles that trace the eye movement, so that in addition to correlating the scores to the parameters, we would be able to objectively determine which body parts of the horse were observed by each expert.

4.4. General Limitations of the Study

4.4.1. Sample Size and Statistical Modelling

One aspect of this study to consider is the relatively small sample size, equal or larger than many previous studies [12,13,21,28], but smaller than others focusing on breeding value estimation [16]. A sample size of 24 horses is limited considering the amount of parameters that were included in the LMEs, which had consequences on the models and their outcome.

The effect sizes η^2 depend on the model and should normally be compared to values estimated for similar studies [33]; however, none of the previously cited studies used this statistic. Suggestive thresholds of small effects $\eta^2 = 0.01$, medium effects at $\eta^2 = 0.04$ and large effects at $\eta^2 = 0.14$ were mentioned in [33]. The parameters we retained therefore showed medium to large effects on SL and OTD despite the relatively small sample sizes.

Considering the small standard deviation of 3 cm, we consider the effects of size negligible on this sample, while being aware that height at withers or limb length should be included in investigations of linear parameters such as SL and OTD in samples with a higher variation in horse size.

4.4.2. Treadmill Locomotion Versus Overground Locomotion

This study was performed on the treadmill, in a controlled environment to assess the effects of speed on our kinematic parameters. At walk, ridden French Saddle horses took longer strides and had consequently lower stride frequencies than when ridden on a sand track, at comparable speeds [36]. At the trot, Buchner et al [37] previously showed that the relative stance duration of the forelimbs was longer, and Ret_{max}_front and Ret_{max}_hind angles were larger compared to overground locomotion. The same phenomenon is likely affecting the walk, although to our knowledge, there have been no similar studies investigating the effects of treadmill locomotion on the limb kinematics at walk in unridden horses.

5. Conclusions

In this study, we confirmed that at walk, ground coverage is related both to stride length and the movement of the forehand, quantifiable with OMC measuring the maximal retraction angle of the forelimbs, the range of yaw of the forehand, the maximal fetlock hyperextension of the forelimb and the maximal forelimb protraction angle. Using IMUs, stride length (and therefore ground coverage) can be quantified measuring the maximum and range of motion of pelvis pitch, and the maximal protraction angle of the metacarpus. Over-tracking at the walk can be quantified with the range of pelvis pitch, the maximal protraction of the hind limbs and the maximum fetlock hyperextension of the hind limbs using OMC, or with the maximum and range of pelvis pitch, maximum metatarsal protraction angle and the contralateral pro-retraction angles of the metatarsus if only IMUs are available. Future field studies need to determine how these parameters could be measured during breeding evaluations such as the field test in the FM and other breeds.

Author contributions

Conceptualisation: AIG, AMC, MN, MAW; Methodology: AIG, EHH, FMSB, AMC, MAW, MN; Software: EHH, FMSB, MAW; Formal analysis: AIG, EHH; Investigation: AIG, FMSB, AMC, MAW; Resources: AIG, MAW, MN; Data curation: AIG, MAW, EHH, FMSB; Writing – original draft preparation: AIG; Writing – review and editing: AIG, MAW, MN, AMC; Visualisation: AIG; Supervision: MAW, MN; Project administration: AIG, MAW, MN; Funding acquisition: AIG, MN, MAW.

Acknowledgments

We are gratefully acknowledging the assistance of Ugo Maninchedda, Nina Waldern, Thomas Wiestner and Christian Gerber during data collection. We thank all the judges participating in the experiments for their valuable time and insights, and private horse owners for providing the stallions for the study's purpose. We would also like to thank Dr Pierre-André Poncet, former director of the Swiss national stud farm and main instructor of the FM experts, for his valuable comments with regard to the formal definitions of traits.

Financial disclosure

This study was funded by the Swiss Federal Office for Agriculture (FOAG) under contract numbers 625000469 and 627001325.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.jevs.2022.104024.

References

[1] Koenen E, Aldridge L, Philipsson J. An overview of breeding objectives for warmblood sport horses. Livestock Production Science 2004;88(1-2):77–84.
[2] Gmel AI, Gmel G, von Niederhäusern R, Weishaupt MA, Neuditschko M. Should we agree to disagree? An evaluation of the inter-rater reliability of gait quality traits in franches-montagnes stallions. J Equine Vet Sci 2020;88:1–5 102932.
[3] Hawson L, McLean A, McGreevy P. Variability of scores in the 2008 Olympic dressage competition and implications for horse training and welfare. J Vet Behav 2010;5(4):170–6.
[4] Stachurska A, Bartyzel K. Judging dressage competitions in the view of improving horse performance assessment. Acta Agriculturae Scandinavica, Section A-Animal Science 2011;61(2):92–102.
[5] Stachurska A, Pięta M, Niewczas J, Markowski W. The freestyle dressage competition as a test of the horse's performance. Equine Comp Exercise Physiol 2006;3(2):93–100.
[6] Moore G, Nankervis K, Dumbell L, Douglas J. A comparison of Piaffe in competition with the Fei definition of the movement. Equine Vet J 2014;46 44–44.
[7] Fuller CJ, Bladon BM, Driver AJ, Barr AR. The intra-and inter-assessor reliability of measurement of functional outcome by lameness scoring in horses. Vet J 2006;171(2):281–6.
[8] Hewetson M, Christley R, Hunt I, Voute L. Investigations of the reliability of observational gait analysis for the assessment of lameness in horses. Vet Rec 2006;158(25):852–8.
[9] Hammarberg M, Egenvall A, Pfau T, Rhodin M. Rater agreement of visual lameness assessment in horses during lungeing. Equine Vet J 2016;48(1):78–82.
[10] Bragança FS, Brommer H, van den Belt A, Maree J, van Weeren P, van Oldruitenborgh-Oosterbaan MS. Subjective and objective evaluations of horses for fit-to-compete or unfit-to-compete judgement. Vet J 2020 105454.
[11] Holmström M, Fredricson I, Drevemo S. Biokinematic differences between riding horses judged as good and poor at the trot. Equine Vet J 1994;26(S17):51–6.
[12] Back W, Barneveld A, Bruin G, Schamhardt H, Hartman W. Kinematic detection of superior gait quality in young trotting warmbloods. Vet Q 1994;16(sup2):91–6.
[13] Back W, Schamhardt H, Barneveld A. Are kinematics of the walk related to the locomotion of a warmblood horse at the trot? Vet Q 1996;18(sup2):79–84.
[14] Biau S, Barrey E. Relationship between stride characteristics and scores in dressage tests. Pferdeheilkunde 2004;20(2):140–4.
[15] Deuel NR, Park J-J. The gait patterns of Olympic dressage horses. International Journal of Sport Biomechanics 1990;6(2):198–226.

[16] Molina A, Valera M, Galisteo A, Vivo J, Gómez M, Rodero A, Agüera E. Genetic parameters of biokinematic variables at walk in the Spanish Purebred (Andalusian) horse using experimental treadmill records. Livestock Science 2008;116(1-3):137–45.

[17] Franches-Montagnes Breeding Association: Programme d'élevage [breeding programme]. 2020. https://www.fm-ch.ch/sites/default/files/content/elevage/reglements_et_directives/programme_delevage_fsfm_f_en_vigueur_des_2020_modif_en_rouge.pdf [accessed 07.04.22].

[18] Skyrme D. MSc Thesis. A comparison of visual scan patterns for an elite judge, elite coach and novice judge in dressage. Chester, UK: University of Chester; 2013.

[19] Wolframm I, Schiffers H, Wallenborn A. Visual attention in Grand Prix dressage judges. Journal of Veterinary Behavior: Clinical Applications and Research 2013;2(8):e25.

[20] Smith J. The Spanish walk & sport. SA Horseman 2009;4(4):44–5.

[21] Clayton HM. Comparison of the stride kinematics of the collected, medium, and extended walks in horses. Am J Vet Res 1995;56(7):849–52.

[22] Davidson EJ. Lameness evaluation of the athletic horse. Veterinary Clinics: Equine Practice 2018;34(2):181–91.

[23] Gmel AI, Gmel G, Weishaupt MA, Neuditschko M. Gait quality scoring data of Franches-Montagnes stallions at walk and trot on a treadmill by experts of the breed and their reliability. Data in Brief 2022:1–11 108123.

[24] Weishaupt MA, Hogg HP, Wiestner T, Denoth J, Stüssi E, Auer JA. Instrumented treadmill for measuring vertical ground reaction forces in horses. Am J Vet Res 2002;63(4):520–7.

[25] Weishaupt MA, Byström A, von Peinen K, Wiestner T, Meyer H, Waldern N, Johnston C, Van Weeren R, Roepstorff L. Kinetics and kinematics of the passage. Equine Vet J 2009;41(3):263–7.

[26] Clayton HM, Hobbs S-J. The role of biomechanical analysis of horse and rider in equitation science. Applied Animal Behaviour Science 2017;190:123–32.

[27] Weishaupt M, Hogg H, Auer J, Wiestner T. Velocity-dependent changes of time, force and spatial parameters in Warmblood horses walking and trotting on a treadmill. Equine Vet J 2010;42:530–7.

[28] Clayton HM. Comparison of the stride kinematics of the collected, working, medium and extended trot in horses. Equine Vet J 1994;26(3):230–4.

[29] Gmel AI, Maninchedda U, Cruz A, Weishaupt M. Video clips of Franches-Montagnes stallions at walk and trot on a treadmill during an incremental speed test. Mendeley Data. Mendeley; 2022. https://data.mendeley.com/datasets/bwfkyvy593/1. [Accessed 22 May 2022].

[30] Kuznetsova A, Brockhoff PB, Christensen RH. lmerTest package: tests in linear mixed effects models. J Stat Softw 2017;82(1):1–26.

[31] Lüdecke D, Ben-Shachar MS, Patil I, Waggoner P, Makowski D. Performance: an R package for assessment, comparison and testing of statistical models. J Open Source Software 2021;6(60):1–8 3139.

[32] Ben-Shachar MS, Lüdecke D, Makowski D. Effectsize: estimation of effect size indices and standardized parameters. J Open Source Software 2020;5(56):1–7 2815.

[33] Lakens D. Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. Front Psychol 2013;4:1–12 863.

[34] Krehbiel TC. Correlation coefficient rule of thumb. Decis Sci J Innov Educ 2004;2(1):97–100.

[35] Darbandi H, Serra Bragança F, Van der Zwaag BJ, Voskamp J, Gmel AI, Haraldsdóttir EH, et al. Using different combinations of body-mounted IMU sensors to estimate speed of horses—a machine learning approach. Sensors 2021;21(3):1–12 798.

[36] Barrey E, Galloux P, Valette J, Auvinet B, Wolter R. Stride characteristics of overground versus treadmill locomotion in the saddle horse. Cells Tissues Organs 1993;146(2-3):90–4.

[37] Buchner H, Savelberg H, Schamhardt H, Barneveld A. Limb movement adaptations in horses with experimentally induced fore-or hindlimb lameness. Equine Vet J 1996;28(1):63–70.