

Kinematic Variability in Running: A Caution Regarding Use of the Spanning Set Measure

Michael Hanlon,¹ Philip Kearney,² and Joan Condell¹

¹University of Ulster; ²University of Chichester

The spanning set technique quantifies intertrial variability as the span between polynomial curves representing upper and lower standard deviation curves of a repeated movement. This study aimed to assess the validity of the spanning set technique in quantifying variability and specifically to determine its sensitivity to variability presented at different phases of a movement cycle. Knee angle data were recorded from a male participant completing 12 overground running trials. Variability was added to each running trial at five different phases of the running stride. Ten variability magnitudes were also used to assess the effect of variability magnitude on the spanning set measure. Variability was quantified in all trials using mean deviation and the spanning set measure. Results of a repeated-measures ANOVA showed significant differences between the spanning set score for trials using different phases of added variability. In contrast, mean deviation values showed no difference related to the phase of added variability. Therefore, the spanning set technique cannot be recommended as a valid measure of intertrial movement variability.

Keywords: spanning set, mean deviation, intertrial variability, movement variability, coefficient of variation

Human movement variability, historically viewed as noise or error in movement patterns, has seen its interpretation evolve recently to a perspective that suggests a functional role of variability (Bartlett et al., 2007). Research in this area has yet to provide a clear indication on the specific directionality of such functional effects, with reports of both positive and negative relationships between variability and performance or health (Bradshaw et al., 2007; Button et al., 2003; Crowther et al., 2008; James et al., 2000). It appears that the effect of variability may result from a complex interaction between the form of variability (e.g., coordination variability, intertrial variability, measurement variability, etc.), performance variable, skill type, and level of performer (Button et al., 2003; Heiderscheidt et al., 2002). Further research is required to address these uncertainties, and such research clearly requires valid and informative measures with which to quantify variability.

The spanning set (SS) is one measure of intertrial variability that has featured in recent research (e.g., Crowther et al. 2008; Kong and Candelaria, 2009). Kurz et al. (2003) originally proposed the SS as an alternative and more sensitive measure than traditional variability

measures such as mean deviation (MD) and coefficient of variation. The mechanics of the SS approach for variability assessment are based on work by Lay (2000) and described in detail by Kurz and Stergiou (2004). In brief, the technique first involves fitting high-order polynomials to the standard deviation (SD) curves of a mean ensemble curve. Seventh-order polynomials were used by both Kurz and Stergiou (2003) and Kurz et al. (2003). The coefficients of each polynomial are then used to define the two vectors of a spanning set between the two SD curves. These two vectors are referred to as \mathbf{u} and \mathbf{v} , which represent the vectors formed by the polynomial coefficients of the upper and lower SD curves, respectively. These polynomial coefficients, from a_0 to a_n (where a is the coefficient and n is the highest polynomial order fitted), differ for each polynomial. The spanning set value is defined as the difference between the two spanning set vectors (see Equation 1). This difference is calculated as the root sum of squared differences between matching coefficient pairs. The greater the spanning set value (y), the greater the variability that is indicated in the mean ensemble curve.

$$y = \|\mathbf{u} - \mathbf{v}\| \quad (1)$$

where y is the magnitude of the spanning set, \mathbf{u} and \mathbf{v} represent the spanning set vectors defined above.

The SS technique has been used to quantify variability in running kinematics between different footwear types (Kurz and Stergiou, 2003) and in walking kinematics between control subjects and patients with peripheral arterial disease (Crowther et al., 2008). While Kurz and Stergiou (2003) reported the SS as a valid measure of

Michael Hanlon (*Corresponding Author*) is with the Sport and Exercise Sciences Research Institute, University of Ulster, Jordanstown, Northern Ireland. Philip Kearney is with the Faculty of Sport, Education and Social Sciences, University of Chichester, Chichester, U.K., and Joan Condell is with the Computer Sciences Research Institute, University of Ulster, Coleraine, Northern Ireland.

stance-phase intertrial variability, Kong and Candelaria (2009) subsequently questioned its use specifically over the full gait cycle in overground barefoot running. They indicated that use of the SS for quantifying variability outside this specific condition may still be appropriate, particularly for stance phase only variability.

Despite these findings both in favor of the SS technique (Kurz et al. 2003; Kurz and Stergiou 2003) and critical of its use in specific circumstances (Kong and Candelaria, 2009), no research has strategically assessed its validity as a measure of intertrial variability in a controlled manner. In reviewing the SS technique, there are indications that the mathematical procedures that underpin it may be overly biased toward variability early in the movement cycle and less sensitive to variability occurring later in the movement cycle. Therefore, the purpose of this study was to assess the validity of the SS measure in a controlled manner using incidences of discrete phase-specific variability added to data from human movement trials.

Method

Subjects and Data Collection

A single-subject design was used to test the validity of the SS technique in quantifying intertrial movement variability. The male subject (age 29 y; height 1.95 m; mass 85.5 kg) was familiarized with the test procedure before providing informed consent to participate in the study. Reflective markers were placed on the subject's right lower extremity as previously described in Kurz and Stergiou (2003) to allow recording of the 2-D sagittal plane knee angle. Marker coordinates were recorded at 500 Hz using an eight-camera Oqus 3-series optoelectronic motion analysis system (Qualisys AB, Gothenburg, Sweden).

After a 5 min warm-up period, the subject completed 12 acceptable runs along a 15 m indoor runway at a range of speeds (2 m/s to 3.5 m/s). Runs were deemed acceptable when the first right foot contact was fully inside an embedded piezoelectric force platform (Kistler, Switzerland) sampling at 1000 Hz and there was no evidence of force platform targeting (visually assessed).

Data Processing

Marker coordinate data were filtered using a Butterworth 4th-order low-pass filter with a cut-off frequency of 14 Hz (Kurz et al. 2003). Initial and final foot contact were determined based on a 5 N force plate threshold. Right knee angle data were then time-normalized to one stride of 101 points (0–100%) using a cubic spline fit with customized software in LabVIEW 8.6 (National Instruments, Texas, USA).

Technique Validation

Due to questions about the phase-related sensitivity of the SS technique, it was assessed using five alternative phase-variability models. Each phase-variability model involved adding simulated variability at a different specific phase of the movement cycle (0–20%, 20–40%, 40–60%, 60–80% or 80–100%). In each phase-variability model, the MD value was kept constant, with just the timing of the variability being altered. These phase-variability models were applied to all twelve running trials and were compared against a constant variability control model. The constant variability control model featured a 2° *SD* at each point (0–100%) of the stride cycle. To assess the effect of variability magnitude on the SS technique, ten different variability magnitudes were also applied in each of the phase-variability models. Table 1 details all 10 variability magnitudes used. Figure 1 illustrates an example of the

Table 1 Variability magnitudes in each of the 10 magnitude models relative to the control model variability. The temporal location of the added variability phase (20% of the cycle) differed for each phase-variability model.

Variability Magnitude	Full Cycle MD (°)	Control	Added Variability	
		Phase (80%) MD (°)	Phase (20%) MD (°) Peak SD (°)	
Control	2	N/A	N/A	N/A
1 +10%	2.2	2	3	4
2 +20%	2.4	2	4	6
3 +30%	2.6	2	5	6
4 +40%	2.8	2	6	8
5 +50%	3.0	2	7	8
6 +60%	3.2	2	8	10
7 +70%	3.4	2	9	10
8 +80%	3.6	2	10	12
9 +90%	3.8	2	11	14
10 +100%	4.0	2	12	14

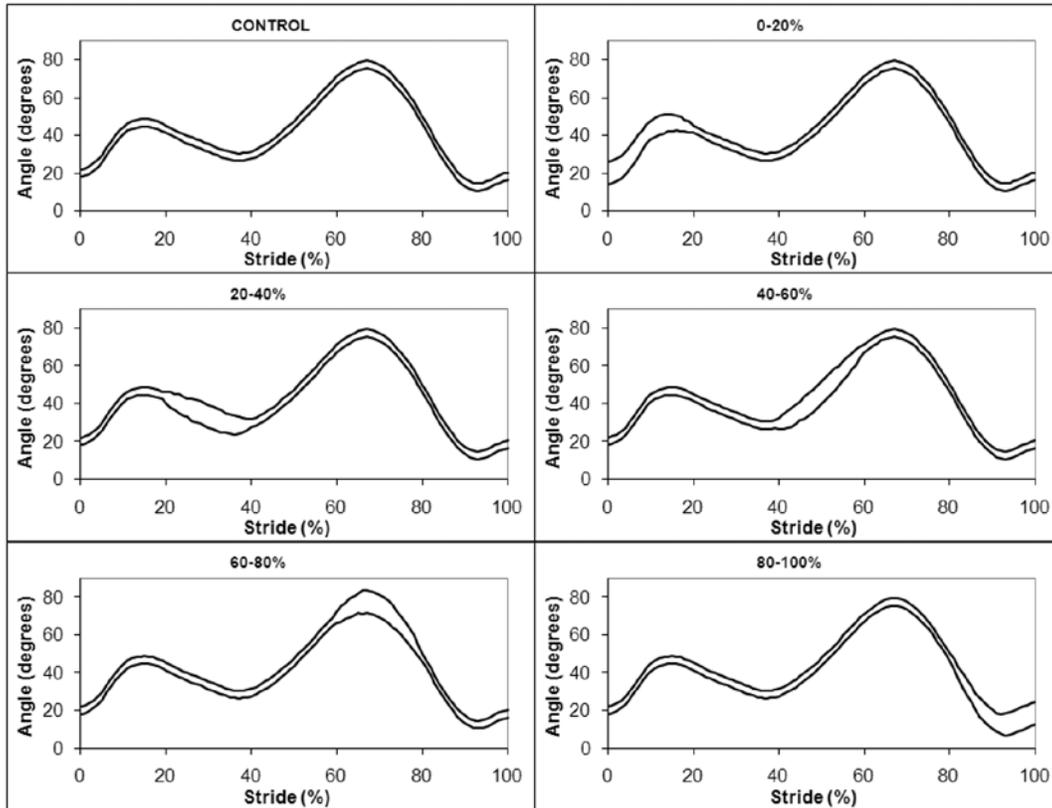


Figure 1 — Control model and five phase-variability models applied to knee angle data from the first running trial. In each of the five phase-variability models shown, the average variability magnitude across the full stride is 30% greater (based on MD) than in the control model.

five phase-variability models as applied to the first running stride using the third variability magnitude (overall MD 30% greater than the control model).

The mean deviation (MD) was calculated for each variability model as the average *SD* across the 101 data points (0–100% of stride). The SS scores were calculated using customized software in LabVIEW as described by Kurz and Stergiou (2004). A 7th-order polynomial was used as part of this technique (fitted using a least squares approach) as recommended by Kurz and Stergiou (2004). This polynomial order was deemed to provide an appropriate fit to the *SD* curves in the current study (all polynomial curves accounted for >99% of the variability in the original *SD* curves, i.e., all R^2 values were greater than 0.99). As with the MD, greater SS values are expected to indicate greater intertrial movement variability.

Statistical Approach

Average SS and MD values were calculated for the 12 running trials in each combination of variability phase ($n = 5$) and variability magnitude ($n = 10$). This provided 10 variability values (each with a different magnitude of variability) for each variability phase. A repeated-measures ANOVA was then used to compare the SS values across each of the five phase-variability models (0–20%,

20–40%, 40–60%, 60–80% or 80–100%). Mauchly's test of sphericity was used to identify violations of the sphericity assumption, and where significant, a Greenhouse-Geisser correction was used (Vincent, 1999). Where a significant overall effect for phase (five levels) was identified, Bonferroni-corrected pairwise comparisons were used to identify between which phases the significant differences existed. MD values (by their nature) did not differ at all across the phase-variability models as only the timing of the added variability differed, therefore the MD values were not assessed statistically. An alpha level of 0.05 was used throughout.

Results

The SS variability scores differed significantly ($p < .001$) depending on which phase of the stride cycle variability was added. Pairwise comparisons indicated that there was a significant difference for SS variability scores between each of the five phase-variability models. The MD and SS scores for each combination of variability phase and magnitude are presented in Table 2. As the MD values within each variability magnitude level did not change, regardless of the phase during which variability was added, only one MD value is shown for each variability

magnitude level. Figure 2 illustrates the SS variability scores for each of the five phase-variability models relative to the SS variability score in the control model. The addition of variability during the first 20% of the stride resulted in a mean increase in the SS variability score of $323 \pm 142\%$ compared with control, while addition of the same level of variability (according to MD) during the middle 20% of the stride (40–60%) resulted in a mean decrease in the SS variability score of $29 \pm 14\%$.

Discussion

This study assessed the phase-related sensitivity of the SS variability score in an effort to gauge its validity as a measure of intertrial variability. It is evident from the results that the SS measure provides grossly inconsistent results for overall intertrial variability, with the overall measure of intertrial variability being heavily dependant on which phase of the movement cycle shows greatest

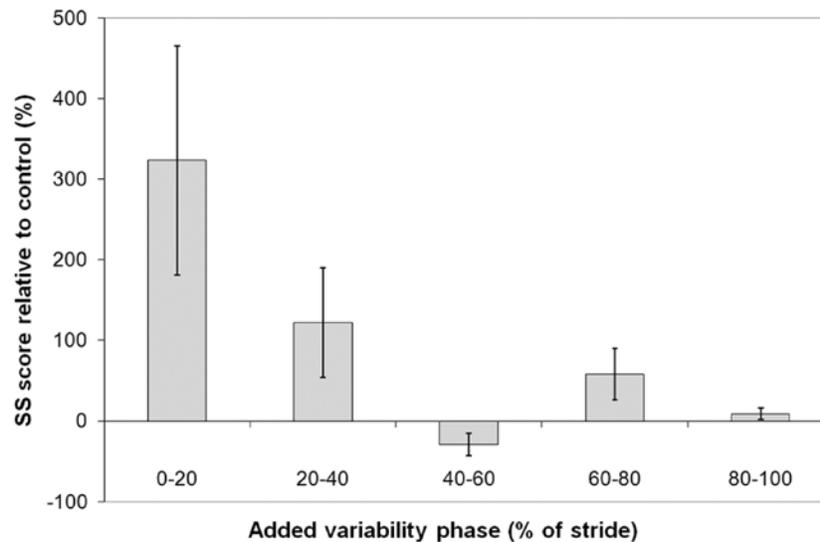


Figure 2 — Average SS variability scores relative to the SS score in the control model (\pm SD) for each added variability phase. For guidance, a value of 100% indicates an SS variability score which is twice that seen in the control model.

Table 2 Average variability scores for MD and SS for each of the 10 variability magnitudes for each phase-variability model. SS values are unitless. The MD and SS scores for the control variability model (for comparison) were 2° and 4° respectively. As MD shows no change related to the phase of variability, single values are used to represent MD for all five phase-variability models in each variability magnitude.

Variability Magnitude	MD (degrees)					SS				
	0–20%	20–40%	40–60%	60–80%	80–100%	0–20%	20–40%	40–60%	60–80%	80–100%
1 +10%			2.2			8.2	4.9	3.5	4.4	4.0
2 +20%			2.4			12.8	5.8	2.9	4.9	4.0
3 +30%			2.6			11.4	6.6	2.6	5.2	4.2
4 +40%			2.8			15.8	7.7	2.3	5.8	4.4
5 +50%			3			14.0	8.3	2.2	6.0	4.3
6 +60%			3.2			18.6	9.4	2.3	6.6	4.6
7 +70%			3.4			17.2	9.9	2.5	6.8	4.1
8 +80%			3.6			21.0	11.3	3.0	7.5	4.7
9 +90%			3.8			26.1	12.2	3.4	7.9	4.7
10 +100%			4			24.2	12.9	3.8	8.2	4.6
Mean			3.1			16.9	8.9	2.8	6.3	4.4
SD			0.61			5.7	2.7	0.6	1.3	0.3

variability. The SS measure appears biased and over-sensitive toward increased variability at the start of the movement cycle (average of 323% increase in scores when variability is added from 0 to 20% of the cycle) while being relatively insensitive to variability occurring at the middle and end of movement patterns (averages of 29% reduction and 9% increase when variability is added from 40 to 60% and 80–100% of the cycle). This is in contrast to MD, a traditional measure of variability, which shows an equal increase in scores regardless of the phase during which variability is added. The consistent decreased estimate of variability by SS when variability is actually added during the 40–60% phase of the cycle highlights the poor relationship between the SS score and intertrial movement variability.

A functional analysis of the SS technique, based on these results, indicates possible causes of the phase-related bias in SS variability scores. As described by Kurz and Stergiou (2004), the SS score uses the calculated differences between coefficient pairs from polynomials mapping the upper and lower *SD* curves. As the first coefficient value is typically the largest, this value has greatest influence on the overall score. However, this first coefficient value also indicates the intercept value of the *SD* curve (i.e., the *SD* value at the start of the movement cycle), hence the particular bias toward early-phase variability. These findings provide an explanation for previous studies reporting unexpected results using the SS technique. Both Kong and Candelaria (2009) and Crowther et al. (2008) reported SS results which did not follow the trend of other variability measures (MD and CV). In each of these studies, the phases showing greatest variability were in the middle and latter stages of the movement cycle, hence the null findings with the SS measure. Both studies, however, concluded that the unexpected SS results were possibly just the result of using a full stride cycle instead of just the stance phase, rather than a result of fundamental issues with the SS technique itself.

Any measure of intertrial variability should be equally sensitive to increases in variability at all phases of the movement cycle. The fact that the SS measure is unequally and somewhat randomly weighted toward different phases of variability means that it is unsuitable for comparisons between subjects or between conditions (as variability could occur during any movement phase). These results prompt reinterpretation of the findings of previous studies supporting the SS technique. For example, the initial claim by Kurz et al. (2003) that the SS technique offered a more sensitive measure of movement variability between shod and barefoot running would appear to be solely the result of increased variability at the start of the movement cycle in barefoot versus shod conditions, rather than increased variability throughout the complete gait cycle. The suggested sensitivity of the SS technique is, therefore, an artifact of the calculation, which only holds true when variability is increased during the initial stage of a movement cycle. This increased sensitivity of the SS technique to early phase variability could be viewed as one positive aspect of the technique as MD

(calculated over a complete stride cycle) can be somewhat insensitive to variability occurring over a short period (e.g., initial loading response). However, research studies aiming to assess intertrial variability changes at these specific phases can address this by using the average *SD* during these specific phases (rather than using the SS) to ensure greater sensitivity is afforded to phases of interest.

In considering alternatives to the SS, the MD appears to be the most appropriate due to its clear, consistent and easily comprehensible link with variability. However, a strong note of caution should be issued in relation to another traditional measure of intertrial variability which is derived from the MD, i.e., the coefficient of variation (CV). While viewed as a useful way of normalizing the *SD* or MD so that variability can be compared between different conditions, individuals and variables, as pointed out by Mullineaux et al. (2001), the inclusion of the mean as the denominator can lead to imbalances between CV values and absolute *SD* values (e.g., when the mean is close to zero).

Finally, it should also be pointed out that assessing intertrial variability can be affected by variations in timing as well as amplitude. Subtle variations in timing (possibly caused by altered speeds) can lead to inflated amplitude variability estimates (Chau et al., 2005). This can be somewhat addressed using curve registration techniques (Sadeghi et al., 2000), however such techniques are not yet widely used. Variability measured using MD without curve registration cannot currently differentiate between temporal and amplitude variability. Further research is required to address the issue of temporal variability in the quantification of intertrial movement variability.

References

- Bartlett, R., Wheat, J., & Robins, M. (2007). Is movement variability important for sports biomechanists? *Sports Biomechanics*, 6, 224–243.
- Bradshaw, E.J., Maulder, P.S., & Keogh, K.L. (2007). Biological movement variability during the sprint start: Performance enhancement or hindrance? *Sports Biomechanics*, 6, 246–260.
- Button, C., Macleod, M., Sanders, R., & Coleman, S. (2003). Examining movement variability in the basketball free-throw action at different skill levels. *Research Quarterly for Exercise and Sport*, 74, 257–269.
- Chau, T., Young, S., & Redekop, S. (2005). Managing variability in the summary and comparison of gait data. *Journal of Neuroengineering and Rehabilitation*, 2, 22.
- Crowther, R.G., Spinks, W.L., Leicht, A.S., Quigley, F., & Gollidge, J. (2008). Lower limb movement variability in patients with peripheral arterial disease. *Clinical Biomechanics (Bristol, Avon)*, 23, 1080–1085.
- Heiderscheit, B.C., Hamill, J., & van Emmerik, E.A. (2002). Variability of stride characteristics and joint coordination among individuals with unilateral patellofemoral pain. *Journal of Applied Biomechanics*, 18, 110–121.
- James, C.R., Dufek, J.S., & Bates, B.T. (2000). Effects of injury proneness and task difficulty on joint kinetic variability. *Medicine and Science in Sports and Exercise*, 32, 1833–1844.

- Kong, P.W., & Candelaria, N.G. (2009). Spanning set may not be appropriate for measuring knee kinematic variability for the entire gait cycle during over-ground barefoot running. *Journal of Applied Biomechanics*, *25*, 258–264.
- Kurz, M.J., & Stergiou, N. (2003). The spanning set indicates that variability during the stance period of running is affected by footwear. *Gait & Posture*, *17*, 132–135.
- Kurz, M.J., & Stergiou, N. (2004). Mathematical measures of coordination and variability in gait patterns. In N. Stergiou (Ed.), *Innovative Analysis of Human Movement* (pp. 163–186). Champaign, IL: Human Kinetics.
- Kurz, M.J., Stergiou, N., & Blanke, D. (2003). Spanning set defines variability in locomotive patterns. *Medical & Biological Engineering & Computing*, *41*, 211–214.
- Lay, D.C. (2000). *Linear algebra and its applications*. New York: Addison-Wesley.
- Mullineaux, D.R., Bartlett, R.M., & Bennett, S. (2001). Research design and statistics in biomechanics and motor control. *Journal of Sports Sciences*, *19*, 739–760.
- Sadeghi, H., Allard, P., Shafie, K., Mathieu, P., Sadeghi, S., Prince, F., et al. (2000). Reduction of gait variability using curve registration. *Gait & Posture*, *12*, 257–264.
- Vincent, W. (1999). *Statistics in kinesiology* (3rd ed.). Champaign, IL: Human Kinetics.

Copyright of Journal of Applied Biomechanics is the property of Human Kinetics Publishers, Inc. and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.