Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

GNSS-Based Latent Performance Dimensions of Race Performance Across Different Wind Conditions in Formula Kite

Javier Riscart-López
Universidad de Cádiz
Luka Pezelj
University of Split
José V. Gutiérrez-Manzanedo
Universidad de Cádiz
Israel Caraballo Vidal
israel.caraballo@uca.es

Universidad de Cádiz
Carlos Manuel Pérez-Pérez
Universidad de Cádiz

Research Article

Keywords: Formula Kite, Sailing performance, GNSS, Principal Component Analysis, Principal Component Regression, Wind conditions

Posted Date: November 24th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10859080/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Background

Formula Kite is an Olympic sailing discipline in which race performance depends on the interaction of multiple technical and tactical factors. Although Global Navigation Satellite System (GNSS)-derived performance variables are widely used in sailing performance analysis, the multidimensional structure underlying these variables and their contribution to race performance under different wind conditions remain poorly understood. This study aimed to identify the latent performance dimensions derived from GNSS-based variables and determine their contribution to race time across different wind speed categories.

Methods

GNSS data were retrospectively obtained from seven international Formula Kite competitions. GNSS-derived performance variables describing the start, sailing trajectory, Velocity Made Good (VMG), manoeuvres, and race characteristics were obtained from the SAP Sailing tracking system. Principal Component Analysis (PCA) was applied to identify latent performance dimensions, and Principal Component Regression (PCR) was subsequently used to determine their contribution to race time under very light, light, moderate, and strong wind conditions.

Results

PCA identified five orthogonal performance dimensions representing Sailing trajectory, Velocity Made Good (VMG), Starting speed, Starting position, and Broad reach performance. PCR demonstrated that Sailing trajectory was the primary determinant of race time across all wind speed categories, whereas the contribution of the remaining performance dimensions varied according to wind conditions. These findings indicate that Formula Kite performance is determined by the interaction of multiple technical and tactical dimensions rather than isolated GNSS-derived variables.

Conclusion

Formula Kite performance can be explained by a limited number of latent performance dimensions derived from GNSS-based variables. A multidimensional analytical approach based on PCA and PCR provides a practical framework for understanding race performance, supporting performance monitoring, and guiding training priorities according to prevailing wind conditions.

Background

Formula Kite is an Olympic sailing discipline characterized by extreme velocities, dynamic hydrofoiling technology [1], and high-speed tactical interactions [2]. Unlike traditional displacement hulls, foiling kitesurfers operate above the water surface, substantially reducing hydrodynamic drag and reaching speeds exceeding 30–40 knots [1]. Consequently, performance in high-performance sailing involves complex interactions between physical, physiological, technical, perceptual-motor, biomechanical, and tactical demands [2, 3]. In this highly dynamic environment, small variations in trajectory or board control can lead to substantial performance discrepancies [1], making the precise identification and quantification of race determinants essential for athletic preparation.

Over the past decade, wearable Global Navigation Satellite System (GNSS) tracking devices have transformed performance analysis across team sports [4] and sailing disciplines [5]. In competitive sailing, tracking systems operating at sampling frequencies of 5 Hz to 10 Hz provide valid and reliable metrics [4] for evaluating speed over ground (SOG) [6], Velocity Made Good (VMG) [7], course trajectories [5], and spatial maneuvering [8]. Previous research across Olympic and international sailing classes—such as the RS:X class [9], Laser class [6], Optimist class [8], and the 2.4mR class [10]—has established the utility of GNSS-derived metrics for characterizing race execution and differentiating performance levels. However, most empirical studies have evaluated GNSS-derived metrics as isolated, independent variables [6, 9].

Evaluating GNSS variables in isolation presents significant statistical and conceptual limitations [11]. In high-performance sailing, metrics such as distance sailed, average VMG, and starting speed exhibit high intercorrelations, introducing substantial multicollinearity when included in conventional linear regression models [1]. Furthermore, single-variable analyses fail to capture the multidimensional nature of sailing performance [9], where tactical execution during the start, upwind leg efficiency, and downwind speed simultaneously interact to determine total race time [8]. A comprehensive analytical framework must therefore account for the latent structure of these interconnected variables to reduce redundancy while retaining the relevant information contained in the dataset [11].

From the ecological dynamics perspective, sailing performance is governed by continuous interactions between athlete capabilities, task demands, and environmental constraints [12]. Wind speed acts as a primary environmental constraint, forcing sailors to continuously adapt their visual search behaviors [13], board control, and tactical decision-making [14, 15] according to task constraints [16]. The physiological demands of sailing further modulate an athlete's ability to maintain optimal performance strategies across different race phases [17]. According to the Royal Yachting Association guidelines, wind conditions can be categorized into distinct operational speed regimes [18]. Previous research suggests that the relative importance of tactical positioning, starting efficiency, trajectory optimization, and sailing execution may vary according to wind conditions [1, 6]. Despite this theoretical grounding, how latent performance dimensions contribute to race time across varying wind speed categories remains unexplored in elite Formula Kite racing.

To address these limitations, advanced dimensionality reduction techniques are required. Principal Component Analysis (PCA) enables the transformation of high-dimensional, highly correlated GNSS datasets into a smaller set of uncorrelated, orthogonal components representing latent performance dimensions [11]. When combined with Principal Component Regression (PCR), this approach reduces multicollinearity while retaining the information captured by the extracted orthogonal components [11]. Therefore, the aims of this study were: (1) to identify the latent performance dimensions underlying GNSS-derived metrics in elite Formula Kite sailors across international competitions, and (2) to determine the relative contribution of these latent dimensions to total race time across four distinct wind speed categories (very light, light, moderate, and strong wind) [18] using Principal Component Regression (PCR). It was hypothesized that GNSS-derived variables would condense into a small set of orthogonal performance dimensions and that the relative contribution of these dimensions to race time would vary according to wind speed category.

Methods

Participants

The study included 140 elite Formula Kite sailors (14 women) who competed in seven international regattas held from 2016 to 2023 (Table 1). In the regattas that included both male and female athletes, competitors raced simultaneously on the same course under identical wind conditions. Consequently, performance data were analysed jointly without sex-based stratification. Complete GNSS data were available for all analysed competitors in six of the seven regattas. The only exception was the 2023 Semaine Olympique Française, in which 92 athletes were registered, but complete GNSS data were available for only 27 competitors, who were consequently included in the analysis.

Table 1
Characteristics of the regattas and athlete participations included in the study.

Regatta	Year	Male	Female	Total analysed
Sailing World Cup Final (Melbourne, Australia)	2016	16	4	20
World Cup Series (Santander, Spain)	2017	16	2	18
World Cup Series (Hyères, France)	2017	16	4	20
Hempel World Cup Series Final (Marseille, France)	2019	16	4	20
Youth Sailing World Championships (Mussanah, Oman)	2021	16	0	16
Allianz Youth Sailing World Championships (The Hague, The Netherlands)	2022	19	0	19
Semaine Olympique Française (Hyères, France)*	2023	27	0	27
Total		126	14	140

*A total of 92 athletes were registered for the 2023 Semaine Olympique Française; however, complete GNSS data were available for only 27 competitors, who were therefore included in the analysis.

Regattas

Data were retrospectively collected from seven international Formula Kite regattas held from 2016 to 2023. These competitions comprised Sailing World Cup events, World Cup Finals, Youth Sailing World Championships, and the Semaine Olympique Française, representing the highest level of international Formula Kite competition during the study period. All races were conducted according to the Racing Rules of Sailing and the corresponding event Sailing Instructions. The racecourse adopted in all regattas consisted of six legs, including two upwind, two downwind, and two reaching legs, from which GNSS-derived performance variables were extracted for each race leg. Race tracking data were obtained from the SAP Sailing® platform [19], which records competitors' trajectories using TracTrac® GPS devices operating at a sampling frequency of 5 Hz, a system whose reliability and validity for high-performance sailing and dynamic trajectory tracking have been established in previous elite sailing research [5, 20]. Wind speed for each race was obtained from the official measurements provided by the race committee. Based on the average wind speed recorded during each race, races were classified according to the Royal Yachting Association (RYA) into four categories: very light wind (5 to ≤ 8 knots), light wind (> 8 to ≤ 12 knots), moderate wind (> 12 to ≤ 15 knots), and strong wind (> 15 knots) [18]. This classification was used in all subsequent analyses.

Outcome and predictor variables

Race time (s) was defined as the outcome variable. Twenty-two GNSS-derived performance variables describing the start and race execution were initially considered as predictor variables. These variables comprised start-related metrics and leg-specific performance indicators obtained throughout the racecourse. Their operational definitions, units of measurement, and data acquisition procedures are presented in Table 2.

Table 2
GNSS-derived predictor variables included in the study

Race phase	Variable	Unit	Description	Data acquisition
	Distance to the starting line (5 s before the start)	**m**	**Distance between the boat and the starting line five seconds before the starting signal.**	**Calculated from GNSS coordinates recorded 5 s before the start.**
	Speed over ground (5 s before the start)	knots	Boat speed over ground measured five seconds before the starting signal.	Obtained directly from the GNSS receiver.
	Distance to the starting line (at the start)	m	Distance between the boat and the starting line at the starting signal.	Calculated from GNSS coordinates at the starting signal.
Starting sequence	Start delay	s	Time difference between the starting signal and the moment the boat crossed the starting line (0 s indicates a perfect start).	Calculated from GNSS timestamps.
	Speed over ground (at the start)	knots	Boat speed over ground at the moment of crossing the starting line.	Obtained directly from the GNSS receiver.
	Speed over ground (immediately after the start)	knots	Boat speed over ground immediately after crossing the starting line.	Obtained directly from the GNSS receiver.
	Distance from the starboard end of the starting line	m	Lateral distance between the boat and the starboard end of the starting line when crossing the line.	Calculated from GNSS coordinates.
	Distance travelled	m	Total distance sailed during the first upwind leg.	Calculated from the recorded GNSS trajectory.
Upwind 1	Velocity Made Good (VMG)	knots	Mean VMG achieved during the first upwind leg.	Calculated relative to the next mark.
	Number of manoeuvres	n	Total number of tacks performed during the first upwind leg.	Identified from trajectory changes.
	Distance travelled	m	Total distance sailed during the first downwind leg.	Calculated from the recorded GNSS trajectory.
Downwind 1	Velocity Made Good (VMG)	knots	Mean VMG achieved during the first downwind leg.	Calculated relative to the next mark.

Note. *Twenty-two GNSS-derived variables were initially considered. Following the variable selection procedure described in Statistical analysis Section, sixteen variables were retained for the final Principal Component Analysis and subsequent Principal Component Regression.*

Race phase	Variable	Unit	Description	Data acquisition
	Distance to the starting line (5 s before the start)	**m**	**Distance between the boat and the starting line five seconds before the starting signal.**	**Calculated from GNSS coordinates recorded 5 s before the start.**
	Number of manoeuvres	n	Total number of gybes performed during the first downwind leg.	Identified from trajectory changes.
	Distance travelled	m	Total distance sailed during the second upwind leg.	Calculated from the recorded GNSS trajectory.
Upwind 2	Velocity Made Good (VMG)	knots	Mean VMG achieved during the second upwind leg.	Calculated relative to the next mark.
	Number of manoeuvres	n	Total number of tacks performed during the second upwind leg.	Identified from trajectory changes.
	Distance travelled	m	Total distance sailed during the second downwind leg.	Calculated from the recorded GNSS trajectory.
Downwind 2	Velocity Made Good (VMG)	knots	Mean VMG achieved during the second downwind leg.	Calculated relative to the next mark.
	Number of manoeuvres	n	Total number of gybes performed during the second downwind leg.	Identified from trajectory changes.
	Distance travelled	m	Total distance sailed during the broad reach leg.	Calculated from the recorded GNSS trajectory.
Broad Reach	Velocity Made Good (VMG)	knots	Mean VMG achieved during the broad reach leg.	Calculated relative to the next mark.
	Number of manoeuvres	n	Total number of manoeuvres performed during the broad reach leg.	Identified from trajectory changes.

Note. *Twenty-two GNSS-derived variables were initially considered. Following the variable selection procedure described in Statistical analysis Section, sixteen variables were retained for the final Principal Component Analysis and subsequent Principal Component Regression.*

Statistical analysis

All statistical analyses were performed using IBM SPSS Statistics (Version 29.0; IBM Corp., Armonk, NY, USA). Descriptive statistics were calculated for all study variables and are presented as mean ± standard deviation (SD), and statistical significance was established at $p < 0.05$.

Principal Component Analysis

The statistical analysis was designed to reduce the dimensionality of the GNSS-derived performance variables and develop predictive models of race time based on the extracted latent components.

Because several variables described related aspects of sailing performance, their simultaneous inclusion in a regression model could introduce redundancy and multicollinearity. Therefore, a Principal Component Analysis (PCA) was performed to transform the original set of correlated variables into a reduced number of uncorrelated components while retaining most of the information contained in the dataset.

The suitability of the data for PCA was assessed using the Kaiser–Meyer–Olkin (KMO) measure of sampling adequacy and Bartlett's test of sphericity.

Variable selection followed an iterative procedure to obtain a robust and interpretable factorial solution. Initially, variables presenting inadequate individual sampling adequacy in the anti-image correlation matrix (MSA < 0.50) were excluded. PCA was then repeated using the remaining variables, and the resulting solution was evaluated through the communalities of each variable. Variables with communality values below 0.30 were subsequently removed, following established recommendations for exploratory PCA. This process was repeated until all retained variables met the predefined criteria for sampling adequacy and communality.

The final PCA was performed using the Principal Components extraction method based on the correlation matrix. Components with eigenvalues greater than 1.0 were retained according to the Kaiser criterion, and an orthogonal Varimax rotation with Kaiser normalization was selected to maximize the interpretability of the extracted orthogonal components. Standardized component scores were subsequently calculated using the regression method and used in the subsequent predictive analyses.

To ensure consistency across all analyses, the standardized component scores obtained from the final PCA performed on the complete dataset were retained and subsequently used in all Principal Component Regression (PCR) models, including the analyses conducted separately for each wind category. This approach ensured that all regression models were based on an identical latent component structure, thereby allowing direct comparison of the relative contribution of each performance dimension across different wind categories.

Principal Component Regression

A Principal Component Regression (PCR) analysis was performed to identify the latent dimensions associated with race time. Because the original GNSS-derived variables exhibited substantial multicollinearity, regression analyses were performed using the orthogonal component scores derived from the PCA instead of the original variables, thereby preserving the multidimensional structure of sailing performance while avoiding multicollinearity.

The standardized component scores obtained from the final PCA were entered as predictor variables, whereas race time (s) was considered the dependent variable. The Enter method was selected because all retained principal components represented theoretically meaningful and mutually orthogonal dimensions of sailing performance identified by the PCA.

Before interpreting the regression models, the assumptions of linear regression were assessed. Independence of residuals was evaluated using the Durbin–Watson statistic, whereas multicollinearity was examined through Tolerance and the Variance Inflation Factor (VIF). Linearity and homoscedasticity were assessed by visual inspection of standardized residual versus standardized predicted value plots, and residual normality was evaluated using histograms and normal probability (P–P) plots.

Model performance was evaluated using the coefficient of determination (R^2), adjusted coefficient of determination (Adjusted R^2), the standard error of the estimate, standardized regression coefficients (β), and the statistical significance of each predictor.

In addition to the global regression model, independent PCR models were performed for each wind category to investigate whether the relative contribution of the latent performance dimensions differed according to wind intensity.

The overall statistical workflow adopted in the present study is summarized in Fig. 1. The analytical workflow comprised variable screening, dimensionality reduction through Principal Component Analysis (PCA), and subsequent Principal Component Regression (PCR) models performed for the complete dataset and separately for each wind category.

Results

Principal Component Analysis

Following the iterative variable screening procedure, six of the 22 GNSS-derived performance variables were excluded because they did not meet the predefined criteria for sampling adequacy or communality. Consequently, the final Principal Component Analysis was conducted using 16 variables.

The adequacy of the retained variables for PCA was confirmed by a Kaiser–Meyer–Olkin (KMO) value of 0.673 and a significant Bartlett's test of sphericity ($\chi^2 = 29774.35$, $df = 120$, $p < 0.001$), indicating sufficient correlations among the variables to justify factor extraction. Five components with eigenvalues exceeding 1.0 were selected, explaining 78.31% of the total variance (Table 3). The first two components accounted for 44.63% of the total variance, whereas the remaining three components explained an additional 33.68%, resulting in a cumulative explained variance of 78.31%.

Table 3
Variance explained by the retained principal components

Component	Eigenvalue	Variance explained (%)	Cumulative variance (%)
PC1 (Sailing trajectory)	3.803	23.77	23.77
PC2 (VMG)	3.338	20.86	44.63
PC3 (Starting speed)	2.679	16.74	61.37
PC4 (Starting position)	1.558	9.74	71.10
PC5 (Broad reach performance)	1.153	7.21	78.31

Note. KMO = 0.673; Bartlett's test: $\chi^2 = 29774.35$, df = 120, $p < 0.001$; 16 variables retained from the initial 22.

The rotated component matrix revealed a clear and interpretable component structure (Table 4). Component 1 was characterized by distance-related variables recorded during the upwind and downwind legs and was therefore labelled Sailing *trajectory*. Component 2 grouped the Velocity Made Good (VMG) variables measured across the course legs and was named *VMG*. Component 3 comprised the variables describing boat speed during the starting phase and was labelled *Starting speed*. Component 4 represented the athlete's spatial position relative to the start line and was named *Starting position*. Finally, Component 5 was exclusively associated with Broad reach distance and was labelled *Broad reach performance*.

Table 4
Rotated component matrix of the retained GNSS-derived performance variables

Variable	*PC1* *Sailing trajectory*	*PC2* *VMG*	*PC3* *Starting speed*	*PC4* *Starting position*	*PC5* *Broad reach performance*
Upwind 2 distance (m)	0.978				
Downwind 1 distance (m)	0.964				
Downwind 2 distance (m)	0.963				
Upwind 1 distance (m)	0.905				
Downwind 1 VMG (knots)		0.822			
Downwind 2 VMG (knots)		0.799			
Upwind 1 VMG (knots)		0.793			
Upwind 2 VMG (knots)		0.784			
Broad reach VMG (knots)		0.643			
SOG at the start (knots)			0.918		
SOG 5 s before the start (knots)			0.832		
SOG after the start (knots)			0.745		
Starboard distance to the start line crossing (m)			0.519		
Distance to the start line (m)				0.972	
Distance to the start line 5 s before the start (m)				0.970	
Broad reach distance (m)					0.842

Note. Only factor loadings ≥ 0.40 are presented. Principal Component Analysis with Varimax rotation and Kaiser normalization. Rotation converged in five iterations.

To further evaluate the stability of the factorial solution, the PCA was repeated using only observations collected under Very light wind conditions. The resulting solution identified four principal components,

explaining 74.03% of the total variance. Although the fifth component identified in the global analysis was not retained, the main factorial structure remained consistent, preserving the Sailing trajectory, VMG, Starting speed, and Starting position components. These findings indicate that the component structure remained stable when the analysis was repeated under Very light wind conditions, supporting the use of the global PCA solution in the subsequent Principal Component Regression analyses.

Principal Component Regression

To determine the contribution of the five principal components identified by the PCA to race performance, a Principal Component Regression (PCR) analysis was performed using race time as the dependent variable and the five principal component scores as predictor variables. The overall regression model was highly significant [F (5,1863) = 7423.898, $p < 0.001$], explaining 95.2% of the variance in race time (R^2 = 0.952; adjusted R^2 = 0.952). The Durbin–Watson statistic was 1.373 (Table 5).

Consistent with the PCA findings, the regression coefficients indicated that all five retained principal components contributed significantly to the prediction of race time (Table 5). Among them, Sailing trajectory (PC1) showed the largest standardized regression coefficient (β = 0.912, $p < 0.001$), followed by VMG (PC2) (β = −0.314, $p < 0.001$). Broad reach performance (PC5) (β = 0.120, $p < 0.001$), Starting position (PC4) (β = 0.063, $p < 0.001$), and Starting speed (PC3) (β = −0.057, $p < 0.001$) also made significant independent contributions to the prediction of race time.

Table 5
Principal Component Regression predicting race time from the retained principal components

Predictor	B	SE	β	t	p
Constant	588.069	1.043	—	563.810	< 0.001
PC1 (Sailing trajectory)	187.866	1.043	0.912	180.068	< 0.001
PC2 (VMG)	−64.780	1.043	−0.314	−62.091	< 0.001
PC3 (Starting speed)	−11.809	1.043	−0.057	−11.319	< 0.001
PC4 (Starting position)	12.900	1.043	0.063	12.364	< 0.001
PC5 (Broad reach performance)	24.664	1.043	0.120	23.640	< 0.001

Note. Principal Component Regression was performed using the five principal component scores obtained from the Principal Component Analysis as predictor variables. The dependent variable was race time (s). The overall model explained 95.2% of the variance in race time (R^2 = 0.952; adjusted R^2 = 0.952; F(5,1863) = 7423.898, $p < 0.001$). The Durbin–Watson statistic was 1.373. As expected for orthogonal principal components, no multicollinearity was observed among the predictors (Tolerance = 1.000; VIF = 1.000 for all components).

Principal Component Regression according to wind conditions

To investigate whether the contribution of the latent performance dimensions differed according to wind speed category, independent Principal Component Regression (PCR) models were performed for each wind category using the standardized component scores derived from the global Principal Component Analysis (Table 6). All regression models were statistically significant ($p < 0.001$), explaining between 86.3% and 97.1% of the variance in race time. The explanatory power of the models remained consistently high across all wind conditions, with the highest coefficient of determination observed under Moderate wind ($R^2 = 0.971$) and the lowest under Very light wind ($R^2 = 0.863$). Sailing trajectory (PC1) was the strongest predictor of race time across all wind categories, with standardized regression coefficients ranging from $\beta = 0.400$ under Very light wind to $\beta = 0.962$ under Moderate wind. Conversely, VMG (PC2) remained a significant predictor under all wind conditions, showing its greatest contributions under Very light wind ($\beta = -0.590$) and Strong wind ($\beta = -0.575$). The contribution of the starting-related components depended on the wind speed category. Starting speed (PC3) and Starting position (PC4) significantly contributed to race performance under Very light, Light, and Strong wind conditions, whereas neither component was significantly associated with race time under Moderate wind. In contrast, Broad reach performance (PC5) was not significantly associated with race performance under Very light wind, became progressively more influential under Light and Moderate wind conditions, and showed a smaller but significant contribution under Strong wind. Overall, these findings demonstrate that although Sailing trajectory represents the primary determinant of Formula Kite race performance across all wind speed categories, the relative contribution of the remaining latent performance dimensions changes according to the wind speed category.

Table 6
Principal Component Regression models predicting race time according to wind conditions

Predictor	Global	Very light wind	Light wind	Moderate wind	Strong wind
Model fit					
R^2	0.952	0.863	0.951	0.971	0.921
Adjusted R^2	0.952	0.859	0.951	0.970	0.920
Standard error of the estimate (s)	45.09	51.94	42.80	31.32	41.68
F	7423.898	235.620	4039.270	1929.440	749.050
p-value	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
Durbin–Watson	1.373	1.739	1.373	1.521	1.825
Principal components (β)					
PC1 – Sailing trajectory	0.912*	0.400*	0.959*	0.962*	0.921*
PC2 – VMG	−0.314*	−0.590*	−0.315*	−0.343*	−0.575*
PC3 – Starting speed	−0.057*	−0.157*	−0.069*	0.002	−0.039*
PC4 – Starting position	0.063*	0.136*	0.077*	0.013	0.051**
PC5 – Broad reach performance	0.120*	0.008	0.153*	0.196*	0.054**

Note. Principal Component Regression (PCR) models were performed using the standardized component scores obtained from the global Principal Component Analysis. Race time (s) was the dependent variable in all models. Values represent standardized regression coefficients (β). Significant coefficients are shown in bold. *$p < 0.05$; ** $p < 0.01$.

Discussion

The present study identified five latent GNSS-derived performance dimensions associated with race time in elite Formula Kite sailors, corresponding to Sailing trajectory, Velocity Made Good (VMG), Starting speed, Starting position, and Broad reach performance. This multidimensional approach extends the traditional analysis of isolated performance indicators by integrating multiple correlated variables into higher-order constructs, providing a more comprehensive understanding of competitive performance [21]. Sailing trajectory emerged as the most consistent determinant of race performance across the overall dataset, highlighting the importance of optimizing the distance sailed and boat positioning throughout the race, as previously reported in Formula Kite [1], RS:X windsurfing [9], Laser sailing [6], and

49er sailing [20]. In contrast, the relative contribution of VMG, Starting speed, Starting position, and Broad reach performance varied according to wind conditions, supporting previous findings in RS:X windsurfing [7], Olympic windsurfing [9, 22], Optimist sailing [8], and 49er sailing [20], which demonstrated that the importance of technical and tactical performance indicators is strongly influenced by environmental conditions, particularly wind intensity.

Previous studies have primarily examined GNSS-derived performance indicators as independent variables, including boat speed, Velocity Made Good (VMG), sailing distance, manoeuvres, and intermediate race positions in Formula Kite [1], RS:X windsurfing [7], and Olympic Laser sailing [6]. Although these studies have considerably advanced the understanding of the technical and tactical determinants of sailing performance, they have generally evaluated each variable separately, despite the strong interrelationships that exist among race-performance indicators. From a performance analysis perspective, considering isolated variables may not fully capture the complexity of competitive sailing, where performance emerges from the interaction of multiple technical, tactical, and environmental factors [21]. In this context, Principal Component Analysis enabled the reduction of numerous correlated GNSS-derived variables into five interpretable latent performance dimensions, providing a more integrated representation of competitive performance while preserving the multidimensional structure of the original dataset [11]. This multivariate approach provides a more comprehensive framework for identifying the key determinants of performance under different environmental conditions.

The prominent contribution of the Sailing trajectory dimension may be explained by the unique characteristics of Formula Kite racing, where minor deviations from the optimal course can rapidly translate into substantial time losses due to the high speeds achieved by hydrofoiling equipment. Consequently, performance depends not only on maximising boat speed but also on selecting and continuously adapting the most efficient sailing line according to the tactical and environmental constraints encountered during the race. This continuous interaction between technical execution and tactical decision-making has previously been highlighted in Formula Kite, where elite sailors constantly adjust their navigation strategy in response to changing environmental conditions and perceptual information [23]. Similarly, previous studies in kiteboard racing have shown that race performance is closely associated with sailing trajectories and displacement patterns across the course rather than with boat speed alone [2]. From a broader sailing perspective, efficient trajectory selection also reflects the interaction between boat characteristics and hydrodynamic behaviour, as the resistance and dynamic response of high-performance sailing craft are strongly influenced by the sailing line adopted [24]. Moreover, successful trajectory optimisation requires continuous perception–action coupling and tactical adaptation throughout the race, reinforcing the view that sailing performance emerges from the continuous interaction between perception, decision-making, and technical execution rather than from isolated technical actions [25].

Although VMG emerged as one of the latent dimensions associated with race performance, its contribution was consistently lower than that of the Sailing trajectory dimension and varied according to wind conditions. This finding suggests that, in Formula Kite, race success depends not only on achieving

high VMG but also on maintaining an efficient sailing trajectory that integrates tactical positioning, course geometry, and continuous adaptation to changing environmental conditions. Consequently, the Sailing trajectory dimension appears to represent a higher-order latent performance construct, encompassing several technical and tactical processes that cannot be fully explained by speed-related variables alone. This interpretation may explain why the Sailing trajectory dimension accounted for a greater proportion of race-performance variability than VMG in the present study. Nevertheless, the importance of VMG remains evident, as previous research has consistently identified it as one of the most sensitive indicators of competitive performance in Formula Kite [1], RS:X [9, 26], and Laser sailing [6]. However, our results further indicate that the influence of VMG is context-dependent rather than universal, as its contribution changed across wind conditions. These findings are consistent with previous studies in Olympic windsurfing, which have shown that the influence of speed-related performance variables differs across wind conditions and race situations [7, 27].

One of the most relevant findings of the present study was that, although the Sailing trajectory dimension consistently remained the principal determinant of race performance, the contribution of the remaining latent dimensions changed according to wind conditions. This observation suggests that Formula Kite performance should not be interpreted as a fixed combination of technical and tactical factors, but rather as a dynamic system in which the relative importance of each performance dimension depends on the environmental context. Similar conclusions have been reported across Olympic sailing disciplines, where the physical, technical, and tactical demands vary substantially with wind intensity and boat characteristics [3, 28]. In Formula Kite, these changes may be even more pronounced because hydrofoiling enables exceptionally high boat speeds, increasing the influence of subtle tactical decisions and continuous trajectory adjustments as environmental conditions evolve [23]. Consequently, our findings suggest that performance models based on a single performance indicator may not adequately represent the complexity of Formula Kite racing, as the relative contribution of technical and tactical variables shifts across different wind scenarios.

Starting speed and Starting position emerged as secondary latent performance dimensions associated with race performance in the global model, although their contributions varied according to wind conditions. Unlike previous studies, which have generally evaluated starting position, starting speed, or first-leg ranking as independent predictors of race outcome in RS:X windsurfing [7, 26], Olympic Laser sailing [6], and Olympic windsurfing [22], the present findings suggest that starting performance is better interpreted as a multidimensional construct integrating positioning and speed during the opening phase of the race. In the global model, both Starting speed and Starting position contributed to race performance, although their standardized effects were considerably smaller than those of Sailing trajectory and VMG. When the analyses were performed separately according to wind conditions, Starting speed and Starting position jointly contributed under very light, light, and strong wind conditions. Neither component contributed under moderate wind conditions, indicating that the tactical importance of the starting phase depends on the prevailing environmental conditions. This non-linear pattern suggests that the influence of the start cannot be explained solely by increasing or decreasing wind intensity but is likely determined by the interaction between environmental conditions and the tactical

demands of the opening phase of the race. These findings reinforce the view that the effectiveness of the start should not be assessed independently but within the broader competitive context, where the opportunities to gain or maintain tactical advantages differ between wind scenarios. Consequently, starting-related performance should be interpreted as context dependent rather than as a universally dominant determinant of success. From an applied perspective, the use of high-resolution GNSS technology enables coaches and performance analysts to objectively quantify the multiple components of starting performance, including positioning and speed, providing a more comprehensive and objective assessment than conventional observational methods [29].

Broad reach performance was identified as a distinct latent dimension, highlighting the specific contribution of this point of sail to overall Formula Kite race performance. Unlike the remaining principal components, this dimension was exclusively characterised by broad reach distance, suggesting that the way sailors negotiate this transition leg represents a distinct technical aspect of performance rather than simply reflecting overall boat speed or VMG. Broad reach performance contributed significantly under light, moderate, and strong wind conditions, whereas no significant contribution was observed under very light wind conditions, indicating that the influence of this race phase is highly dependent on the environmental context. Similar findings have been reported in Olympic windsurfing, where VMG during beam reach became one of the principal determinants of competitive performance under specific wind conditions [26]. Likewise, research in the Optimist class demonstrated that higher-performing sailors achieved superior VMG during broad reach together with more efficient broad reach trajectories [8]. Although these previous studies analysed broad reach variables individually, the present findings indicate that broad reach performance can also emerge as a distinct latent performance dimension when the interrelationships among GNSS-derived variables are considered. In Formula Kite, where hydrofoiling enables exceptionally high speeds during reaching legs, small differences in trajectory selection and board handling may translate into substantial differences in race time [23]. Consequently, this phase of the race should receive specific attention during both performance analysis and training design rather than being regarded simply as a transition between upwind and downwind legs.

Practical applications

The present findings provide several practical implications for coaches and performance analysts working in Formula Kite. Race performance should be evaluated using a multidimensional approach rather than relying on isolated GNSS-derived variables, as multiple technical and tactical performance dimensions contribute simultaneously to race outcomes. Sailing trajectory emerged as the most influential performance dimension across all wind speed categories, highlighting the importance of optimising tactical decision-making and maintaining efficient sailing trajectories throughout the race. Furthermore, the relative contribution of the remaining performance dimensions varied according to wind conditions, indicating that training priorities and performance analyses should be adapted to the expected environmental scenarios rather than applying a single performance model across all races. The integration of wearable GNSS technology into daily training and competition enables coaches to objectively quantify these multidimensional performance characteristics, providing detailed information

that supports evidence-based decision-making, training monitoring, and athlete performance evaluation [30, 31, 32]. In addition, the combined application of Principal Component Analysis and Principal Component Regression offers a practical analytical framework for reducing the complexity of large GNSS-derived datasets while preserving the most relevant performance information [11]. The resulting multidimensional information can facilitate individualised feedback, training prescription, and performance evaluation in elite Formula Kite sailors [33].

Limitations

Several limitations of the present study should be acknowledged. First, the retrospective design restricted the analyses to the variables available from the SAP Sailing® tracking system, precluding the inclusion of physiological, biomechanical, perceptual-cognitive, and psychological factors that may also contribute to Formula Kite performance. Future research should therefore integrate GNSS-derived performance indicators with physiological, biomechanical, perceptual-cognitive, and psychological measures, as physiological demands have also been shown to vary substantially across different phases of elite sailing competition [34]. Consequently, the latent dimensions identified in the present study should be interpreted as representing the technical and tactical components of performance derived from GNSS data rather than the complete spectrum of determinants influencing race outcomes. Second, although the combined application of Principal Component Analysis and Principal Component Regression proved effective for reducing data complexity and identifying the principal latent performance dimensions, future studies should validate this analytical framework in other Olympic foiling classes and investigate whether integrating physiological, biomechanical, perceptual-cognitive, and environmental variables further improves performance prediction. Third, although female and male sailors competed simultaneously on the same racecourse under identical environmental conditions, the marked imbalance between male and female participants precluded robust sex-specific analyses. Future investigations including more balanced cohorts should therefore examine whether the latent performance dimensions identified in the present study differ according to sex. Finally, although the present study analysed data collected during official international competitions under ecologically valid conditions, the hierarchical structure of repeated race observations within sailors and regattas was not explicitly modelled. Future studies applying hierarchical or mixed-effects statistical approaches across multiple competitive seasons may provide further insight into the stability of these latent performance dimensions and account more explicitly for the nested structure of sailing performance data.

Conclusion

The present study demonstrated that Formula Kite race performance can be explained by a limited number of latent performance dimensions derived from GNSS-based variables. The combined application of Principal Component Analysis and Principal Component Regression provided a multidimensional framework for reducing the complexity of race-performance data while identifying the principal technical and tactical determinants associated with competitive success. Sailing trajectory consistently emerged as the primary determinant of race performance across all wind conditions,

whereas the relative contribution of VMG, Starting speed, Starting position, and Broad reach performance varied according to the prevailing environmental conditions. These findings demonstrate that Formula Kite performance is best understood as the interaction of multiple technical and tactical dimensions rather than isolated performance variables. From an applied perspective, this multidimensional approach provides coaches and performance analysts with an objective framework for monitoring athlete performance, individualising training interventions, and optimising race preparation according to the expected wind conditions. More broadly, the present findings support the growing application of wearable GNSS technologies and multivariate analytical approaches for advancing evidence-based performance analysis in elite sailing.

Abbreviations

GNSS Global Navigation Satellite System

GPS Global Positioning System

SOG Speed Over Ground

VMG Velocity Made Good

PCA Principal Component Analysis

PCR Principal Component Regression

KMO Kaiser–Meyer–Olkin

MSA Measure of Sampling Adequacy

SD Standard Deviation

VIF Variance Inflation Factor

RYA Royal Yachting Association

Declarations

Acknowledgements

The authors would like to thank Luka Pezelj (Faculty of Maritime Studies, University of Split, Split, Croatia) for his valuable collaboration and support within the framework of the SEA-EU Alliance and the cooperation between the University of Cádiz (Spain) and the University of Split (Croatia).

Authors' contributions

Conceptualization, L.P. and I.C.-V.; methodology, L.P., J. V. G. M. and I.C.-V.; software, L.P. and I.C.-V.; validation, L.P. and I.C.-V.; formal analysis, L.P., J.V. G. M. and I.C.-V.; investigation, L.P. and I.C.-V.; resources, L.P. and I.C.-V.; data curation, J.R.-L. and C.M.P.-P.; writing—original draft preparation, J.R.-L., J.V.G.M. and C.M.P.-P.; writing—review and editing, J.R.-L. and C.M.P.-P.; visualization, J.R.-L., C.M.P.-P., L.P. and I.C.-V.; supervision, J.R.-L., C.M.P.-P., L.P. and I.C.-V.; project administration, L.P. and I.C.-V.; funding acquisition, L.P. All authors have read and agreed to the published version of the manuscript.

Funding

This research received no specific funding.

Data availability

The data analysed in this study were derived from historical Formula Kite race-tracking records from seven international regattas held between 2016 and 2023 using the SAP Sailing® race-tracking platform (SAP Sailing Analytics, https://www.sapsailing.com). The regattas and corresponding study samples are identified in Table 1 of the manuscript. The complete historical race-tracking datasets used for the analyses do not have DOI or accession identifiers and their continued public availability through the SAP Sailing® platform cannot be guaranteed. The processed data supporting the findings of this study are available from the corresponding author upon reasonable request.

Ethics approval and consent to participate

Ethical approval was not required because this study was based exclusively on publicly available race tracking data obtained from the SAP Sailing® platform. The study involved no direct interaction with athletes, no intervention, and no collection of private or identifiable personal data. All analyses were conducted using publicly accessible competition records.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

References

1. Caraballo I, Pezelj L, Riscart-López J, Bujalance-Moreno P, Castro-Piñero J. Performance analysis in Olympic sailors of the Formula Kite class using GPS. Sens (Basel). 2021;21(2):574. 10.3390/s21020574.
2. Caimmi G, Semprini G. Heart rate and GPS data analysis of kiteboard course racing during the Italian Championships. Sport Sci Health. 2017;13(1):79–85. 10.1007/s11332-017-0346-9.

3. Spurway N, Legg S, Hale T. Sailing physiology. J Sports Sci. 2007;25(10):1073–5. 10.1080/02640410601165171.
4. Scott MTU, Scott TJ, Kelly VG. The validity and reliability of global positioning systems in team sport: a brief review. J Strength Cond Res. 2016;30(5):1470–90. 10.1519/JSC.0000000000001221.
5. Caraballo I, Conde-Caveda J, Pezelj L, Milavić B, Castro-Piñero J. GNSS applications to assess performance in Olympic sailors: Laser class. Appl Sci. 2021;11(1):264. 10.3390/app11010264.
6. Pan D, Sun K. Analysis of sailing variables and performance of Laser sailors with different rankings under the condition of certain wind speed. Heliyon. 2022;8(12):e11682. 10.1016/j.heliyon.2022.e11682.
7. Anastasiou A, Jones T, Mullan P, Ross E, Howatson G. Descriptive analysis of Olympic class windsurfing competition during the 2017–2018 regatta season. Int J Perform Anal Sport. 2019;19(4):517–29. 10.1080/24748668.2019.1631054.
8. Caraballo I, Pezelj L, Ramos-Álvarez JJ. Analysis of the performance and sailing variables of the Optimist class in a variety of wind conditions. J Funct Morphol Kinesiol. 2024;9(1):18. 10.3390/jfmk9010018.
9. Caraballo I, Riscart-López J, Pezelj L, Castro-Piñero J. Key performance indicators of Olympic windsurfers during a World Cup: RS:X class. J Sports Sci. 2023;41(10):1047–55. 10.1080/02640414.2023.2180854.
10. Caraballo I, Cruz-Leon C, Pérez-Bey A, Gutiérrez-Manzanedo JV. Performance analysis of Paralympic 2.4mR class sailing. J Sports Sci. 2021;39(Suppl 1):109–15. 10.1080/02640414.2021.1917871.
11. Abdi H, Williams LJ. Principal component analysis. Wiley Interdiscip Rev Comput Stat. 2010;2(4):433–59. 10.1002/wics.101.
12. Araújo D, Davids K, Bennett S, Button C, Chapman G. Emergence of sport skills under constraints. In: Williams AM, Hodges NJ, editors. Skill acquisition in sport: research, theory and practice. 2nd ed. London: Routledge; 2006.
13. Pluijms JP, Cañal-Bruland R, Hoozemans MJM, Savelsbergh GJP. Visual search, movement behaviour and boat control during the windward mark rounding in sailing. J Sports Sci. 2015;33(4):398–410. 10.1080/02640414.2014.942683.
14. Araújo D, Rocha L, Davids K. The ecological dynamics of decision-making in sailing. In: Renshaw I, Davids K, Savelsbergh GJP, editors. Motor learning in practice: a constraints-led approach. London: Routledge; 2010. pp. 131–43. 10.4324/9780203888100-18.
15. Dicks M, Button C, Davids K. Examination of gaze behaviours under in situ and video simulation task constraints reveals differences in information pickup for perception and action. Atten Percept Psychophys. 2010;72(3):706–20. 10.3758/APP.72.3.706.
16. Araújo D, Davids K, Passos P. Ecological validity, representative design, and correspondence between experimental task constraints and behavioral setting: comment on Rogers, Kadar, and Costall (2005). Ecol Psychol. 2007;19(1):69–78. 10.1080/10407410709336951

17. Castagna O, Brisswalter J, Lacour JR, Vogiatzis I. Physiological demands of different sailing techniques of the new Olympic windsurfing class. Eur J Appl Physiol. 2008;104(6):1061–7. 10.1007/s00421-008-0863-y.
18. Royal Yachting Association. Olympic classes speed charts [Internet]. Hamble: Royal Yachting Association. 2019 [cited 2026 Aug 5]. Available from: https://www.rya.org.uk/racing/running-racing/speed-charts
19. SAP Sailing. SAP Sailing race tracking database [Internet]. Walldorf: SAP SE. 2019 [cited 2026 Aug 5]. Available from: https://www.sapsailing.com
20. Gutiérrez-Manzanedo JV, Caraballo Vidal I, García-Unanue J, Pezelj L, Ramos-Álvarez JJ. GPS to assess an Olympic regatta: 49er and 49erFX classes. J Navig. 2025;1–11. 10.1017/S0373463325101355.
21. Hughes M, Bartlett R. The use of performance indicators in performance analysis. J Sports Sci. 2002;20(10):739–54. 10.1080/026404102320675602.
22. Chun S, Park J, Kim T, Kim Y. Performance analysis based on GPS data of Olympic class windsurfing. Int J Perform Anal Sport. 2022;22(3):332–42. 10.1080/24748668.2022.2046332.
23. Terrien E, Huet B, Iachkine P, Saury J. Considering perceptual experiences and adaptive actions in performance analysis of elite Formula Kite riders by combining qualitative data and measured key indicators of performance. Int J Sports Sci Coach. 2024;19(1):368–78. 10.1177/17479541231153273.
24. Begovic E, Bertorello C, Mancuso A, Saporito A. Sailing dinghy hydrodynamic resistance by experimental and numerical assessments. Ocean Eng. 2020;214:107458. 10.1016/j.oceaneng.2020.107458.
25. Araújo D, Serpa S. Toma de decisión dinámica en diferentes niveles de expertise en el deporte de vela. Rev Psicol Deporte. 1999;8(1):103–15.
26. Riscart-López J, Pezelj L, Pérez-Pérez CM, Caraballo-Vidal I. Performance and tactical indicators in elite Olympic RS:X sailing across different wind conditions using GPS data. Sensors. 2026;26:3312. 10.3390/s26113312.
27. Hagiwara M, Ishii Y. Analysis of racing factors in windsurfing under light wind. Med Sci Sports Exerc. 2016;48(5 Suppl 1):1040. 10.1249/01.mss.0000488128.92942.46.
28. Bojsen-Møller J, Larsson B, Aagaard P. Physical requirements in Olympic sailing. Eur J Sport Sci. 2015;15(3):220–7. 10.1080/17461391.2014.955130.
29. Semb EE, Stendal E, Dahlhaug K, Steinert M. Multi-sensor system for analysis of maneuver performance in Olympic sailing. Appl Sci. 2025;15:8629. 10.3390/app15158629.
30. Aughey RJ. Applications of GPS technologies to field sports. Int J Sports Physiol Perform. 2011;6(3):295–310. 10.1123/ijspp.6.3.295.
31. Cummins C, Orr R, O'Connor H, West C. Global positioning systems (GPS) and microtechnology sensors in team sports: a systematic review. Sports Med. 2013;43(10):1025–42. 10.1007/s40279-013-0069-2.

32. Winchcombe M, Varley MC, Hulin BT, Weakley JJS. Quantifying and monitoring training load in field sports using GPS and microtechnology: current applications and future directions. Sports Med. 2021;51(11):2301–23. 10.1007/s40279-021-01504-8.
33. Buchheit M. Monitoring training status with HR measures: Do all roads lead to Rome? Front Physiol. 2014;5:73. 10.3389/fphys.2014.00073.
34. Bay J, Bojsen-Møller J, Nordsborg NB. Reliable and sensitive physical testing of elite trapeze sailors. Scand J Med Sci Sports. 2018;28(3):919–27. 10.1111/sms.12993.

Figures

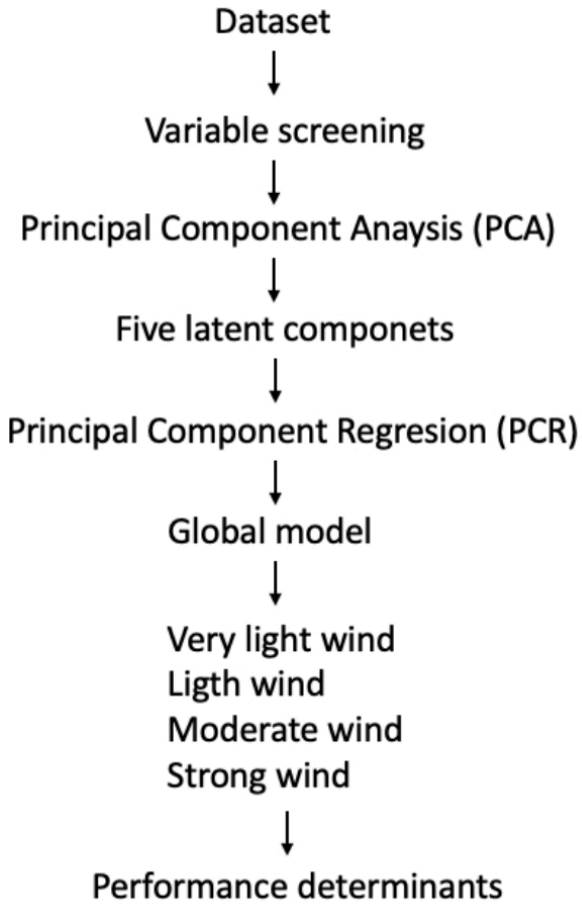

Figure 1

Workflow of the statistical analysis performed to identify the GNSS-derived determinants of Formula Kite race performance.