Universidade do Estado do Rio de Janeiro
Centro de Educação e Humanidades
Instituto de Educação Física e Desportos

Vitor Ayres Principe

Análise de indicadores de desempenho de jogadoras de futebol profissional da elite do futebol brasileiro

Rio de Janeiro
2020

Vitor Ayres Principe

Análise de indicadores de desempenho de jogadoras de futebol profissional da elite do futebol brasileiro

Dissertação apresentada, como requisito parcial para obtenção do título de Mestre, ao Programa de Pós Graduação em Ciências do Exercício e do Esporte, da Universidade do Estado do Rio de Janeiro. Área de concentração: Aspectos Biopsicossociais do Esporte.

Orientador: Prof. Dr. Rodolfo de Alkmim Moreira Nunes

Rio de Janeiro

2020

CATALOGAÇÃO NA FONTE
UERJ/REDE SIRIUS/BIBLIOTECA CEH/B

P957 Principe, Vitor Ayres.
Análise de indicadores de desempenho de jogadoras de futebol profissional da elite do futebol brasileiro / Vitor Ayres Principe. – 2020.
74 f.: il.

Orientador: Rodolfo de Alkmim Moreira Nunes.
Dissertação (mestrado) – Universidade do Estado do Rio de Janeiro, Instituto de Educação Física e Desportos.

1. Futebol feminino – Teses. 2. Aptidão física do atleta – Teses. 3. Inteligência artificial - Aplicações médicas – Teses. I. Nunes, Rodolfo de Alkmim Moreira, 1963-. II. Universidade do Estado do Rio de Janeiro. Instituto de Educação Física e Desportos. III. Título.

CDU 796.332-055.2:613.71

Bibliotecária: Eliane de Almeida Prata. CRB7 4578/94

_____ _____

Assinatura Data

Vitor Ayres Principe

Análise de indicadores de desempenho de jogadoras de futebol profissional da elite do futebol brasileiro

Dissertação apresentada, como requisito parcial para obtenção do título de Mestre, ao Programa de Pós Graduação em Ciências do Exercício e do Esporte, da Universidade do Estado do Rio de Janeiro. Área de concentração: Aspectos Biopsicossociais do Esporte

Aprovada em 20 de fevereiro de 2020.

Banca Examinadora:

_____

Prof. Dr. Rodolfo de Alkmim Moreira Nunes (Orientador)

Instituto de Educação Física e Desportos – UERJ

_____

Prof. Dr. Rodrigo Gomes de Souza Vale

Instituto de Educação Física e Desportos – UERJ

_____

Prof. Dr. Gilson Ramos de Oliveira Filho

Centro Universitário Augusto Motta – UNISUAM

Rio de Janeiro

2020

DEDICATÓRIA

Dedico este trabalho à Deus.

AGRADECIMENTOS

Primeiramente agradeço a Deus por mais uma etapa concluída.

Aos meus pais que sempre me apoiaram em todos os momentos da minha vida.

A minha irmã, meu cunhado e meu sobrinho por estarem presentes em vários momentos.

Ao Prof. Rodolfo Alkmim, meu orientador, que desde o primeiro momento me recebeu me acolheu dentro da universidade.

Ao Prof. Rodrigo Vale pelos conselhos e oportunidades ao longo desta caminhada.

Ao Instituto de Educação Física e Desportos (IEFD), na figura do seu diretor Prof. Carlos Elias e seu vice-diretor Prof. Flávio Chame, pela oportunidade de poder transmitir meus conhecimentos no curso de graduação em Educação Física.

Aos professores e funcionários do Programa de Pós-Graduação em Ciências do Exercício e do Esporte que contribuíram para essa conclusão.

Também agradeço a todos que contribuíram direta e indiretamente com a elaboração desta pesquisa.

“Muitos propósitos há no coração do homem, mas o conselho do Senhor permanecerá.”

(Provérbios 19:21).

RESUMO

PRINCIPE, Vitor Ayres. *Análise de indicadores de desempenho de jogadoras de futebol profissional da elite do futebol brasileiro*. 2020. 74 f. Dissertação (Mestrado em Ciências do Exercício e do Esporte) – Instituto de Educação Física e Desportos, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 2020.

Introdução: Com as ações de promoção do futebol feminino executadas pela CBF existe a expectativa de um aumento significativo no número de times e jogadoras no Brasil. Na literatura, podemos encontrar mudanças drásticas no futebol feminino, representadas pelo aumento da demanda física e pela a capacidade física das atletas. Com o impulso das ferramentas analíticas, oriundas dos avanços científicos, e disponibilidade a grandes quantidades e diversos tipos de dados traz, para ambas as equipes, a necessidade de novas abordagens para o processo análise e compressão, podendo resultar em jogos mais competitivos e em equipes mais preparadas. Desta forma, devemos compreender o atleta de forma global a partir do entendimento da complexidade da modalidade em questão, a fim de identificar indicadores individuais favoráveis à melhora do desempenho esportivo. Objetivo: Verificar e compreender de forma aplicada o uso de tecnologias vestíveis para mensuração das demandas físicas das jogadoras da elite do futebol feminino brasileiro durante as competições. Materiais e métodos: O estudo foi composto por três artigos. O primeiro estudo utilizou-se da metodologia computacional da literatura (CRL) que se utiliza de técnicas de inteligência artificial para observar as lacunas pertinentes da literatura para nortear de forma rápida e objetiva a linha de trabalho a seguir. No segundo, foram selecionados 21 estudos através da revisão sistemática da literatura com base nas recomendações PRISMA, com o objetivo de verificar a aplicabilidade do uso de diferentes sistemas de navegação global por satélite (GNSS) durante partidas e treinamentos de jogadores de futebol. O terceiro estudo buscou-se compreender o comportamento das demandas de carga externa das jogadoras da elite do futebol brasileiro por posição e tempo de jogo a partir dos dados de GPS. Para isso, foram utilizados procedimentos estatísticos e procedimentos de inteligência artificial para modelar o problema estudado. Resultados: O resultado dessa dissertação demostra um aumento significativo em estudos que estão interessados em investigar e poder compreender o comportamento das cargas executadas pelos jogadores de futebol. Neste contexto, o futebol feminino apresenta estudos científicos ainda incipientes, ou seja, o futebol feminino ainda apresenta poucos estudos relacionados a investigar as relações das cargas de treino e jogos. Por ainda estar no início, não foi encontrado nenhum estudo com equipes brasileiras. Desta forma, observa-se que, com as competições que acontecem ao mesmo tempo no calendário do futebol feminino, o controle das cargas passar a ser um desafio e pode trazer um desgaste fisiológico das jogadoras durante as competições. Conclusão: A partir dos resultados encontrados nestes estudos observou-se um decréscimo nas demandas de trabalho externa executadas pelas jogadoras brasileiras de futebol profissional entre suas posições no campo de jogo e entre o 1º e 2º tempo das partidas, sendo necessário um melhor controle das cargas e uma necessidade de melhorar os componentes de condicionamento específicos do futebol à medida que as jogadoras avançam na competição.

Palavras-chaves: Futebol feminino. GPS. Inteligência artificial

ABSTRACT

PRINCIPE, Vitor Ayres. *Performance indicators analysis of woman football elite professional players of brazilian football*. 2020. 74 f. Dissertação (Mestrado em Ciências do Exercício e do Esporte) – Instituto de Educação Física e Desportos, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 2020.

Introduction: With the promotion of women's football implemented by CBF, there is the expectation of a significant increase in the number of teams and players in Brazil. In the literature, we can find drastic changes in women's soccer, represented by the increase in physical demand and the physical capacity of athletes. With the impetus of analytical tools from scientific advances and availability to large amounts and diverse types of data, for both teams, the need for new approaches to the analysis and compression process can result in more competitive games and teams. More prepared. Thus, we must understand the athlete globally from the understanding of the complexity of the sport in question, in order to identify individual indicators favorable to the improvement of sports performance. Objective: To verify and understand the use of wearable technologies to measure the physical demands of Brazilian female elite soccer players during competitions. Materials and methods: The study consisted of three articles. The first study used the computational literature methodology (CRL) that uses artificial intelligence techniques to observe the relevant gaps in the literature to guide quickly and objectively the following line of work. In the second, 21 studies were selected through a systematic literature review based on the PRISMA recommendations, to verify the applicability of the use of different Global Navigation Satellite Systems (GNSS) during soccer players' matches and training. The third study aimed to understand the behavior of the external load demands of the elite Brazilian soccer players by position and to playtime from the GPS data. For this, statistical procedures and artificial intelligence procedures used to model the problem studied. Results: The result of this dissertation demonstrates a significant increase in studies that are interested in investigating and being able to understand the behavior of the loads performed by soccer players. In this context, female soccer presents still incipient scientific studies, that is, female soccer still has few studies related to investigating the relations of training loads and games. Because it is still in its infancy, no study found with Brazilian teams. Thus, it observed that, with the competitions that happen at the same time in the women's soccer calendar, the control of the loads becomes a challenge and can bring the physiological wear of the players during the competitions. Conclusion: From the results found in these studies, there was a decrease in the demands of external work performed by Brazilian professional soccer players between their positions on the playing field and between the 1st and second half of the matches, requiring a better control of the loads. Moreover, a need to improve soccer-specific conditioning components as players advance in the competition.

Keywords: Women's Football. GPS. Artificial Intelligence

SUMÁRIO

INTRODUÇÃO

Na última década houve, um aumento no número de praticantes de futebol feminino, tornando este esporte num dos mais praticados por este público. A FIFA estimou, em 2012, que haveriam mais de 29 milhões de praticantes do sexo feminino mulheres praticando futebol, ou seja, um aumento de 32% relativamente aos dez anos anteriores (FIFA, 2015a).

Em 2015, a Copa do Mundo foi considerada o maior evento esportivo feminino do planeta, com 1 milhão e 350 mil espectadores nos estádios. A cobertura televisiva abrangeu 188 países e com uma audiência de 25 milhões e 400 mil telespectadores na final. Esta foi a maior audiência já registrada em todos os tempos numa partida de futebol feminino, e uma das maiores audiências esportivas nos Estados Unidos (FIFA, 2016)

Existe uma expectativa que haja um crescimento no número de equipes femininas a cada ano no Brasil devido a normativa divulgada em 2017 pela CBF e a CONMBOL (Confederação Sul-Americana de Futebol) onde todos os clubes de futebol participantes do Campeonato Brasileiro, tenham parte dos investimentos direcionados especificamente ao desenvolvimento do futebol feminino. Diante deste fato, a CBF (Confederação Brasileira de Futebol) reformulou as competições profissionais quanto para as categorias de base. Foram criadas a Copa do Brasil Feminina, Série A2 e campeonato brasileiro feminino Sub-20 (RAMOS, 2017).

No futebol feminino tem ocorrido uma mudança drástica nas últimas décadas, conforme citação dos autores Haugen; Tønnessen; Seiler, (2012). Estas demonstram a necessidade de uma elevada capacidade aeróbica para a prática da modalidade, onde são observadas frequências cardíacas médias entre 84% a 86% da frequência cardíaca máxima, em partidas oficiais e distâncias médias percorridas entre 9 e 12 km (HEWITT; NORTON; LYONS, 2014), resultados estes semelhantes as aos encontrados no futebol masculino (BRADLEY, 2010).

Embora o desempenho individual esteja associado à capacidade aeróbica em jogos de futebol (MOHR, KRUSTRUP, & BANGSBO, 2003), as ações de alta intensidade (correr em alta velocidade, saltar, mudar de direção) estão associados aos momentos determinantes do jogo (JULLIEN, 2008).

O crescimento massivo de dados disponíveis para análise oferece várias oportunidades únicas de pesquisa. Portanto, existe uma grande quantidade de dados também que podem ser

coletados acerca dos atletas profissionais, o que faz surgir no futebol a era denominada de *Big Data*. Este conceito é caracterizado pela sua capacidade de acesso a grandes volumes de dados que podem ser úteis para visualização e análise de padrões únicos ou repetidos, relativos a um determinado comportamento (KRISHNAN, 2013; MORGULEV; AZAR; LIDOR, 2018).

Com o acréscimo de volume de dados e o surgimento de técnicas avançadas de análise, que rapidamente são absorvidas pela indústria esportiva, permite que as equipes esportivas possam tirar proveito das mesmas, para o desenvolvimento das equipes (MAXCY; DRAYER, 2014).

Caracterizada como a observação dos objetivos e comportamentos que ocorrem no jogo durante a competição, a análise no esporte vêm variando em sofisticação a partir de dados discretos sobre a atividade de um atleta, numa perspectiva individual, bem como uma visão geral da interação entre os indivíduos (CARLING; WILLIAMS; REILLY, 2005).

A ideia de observar o comportamento humano é historicamente bem estabelecida. Há evidências de que os antigos egípcios usavam hieróglifos para registrar as características do movimento. Os sistemas atuais são bem mais poderosos e complexos e suas utilizações contemporâneas podem ir além da análise dos jogos e rumo à modelagem e predição matemática de futuras competições. Ambos, estilos e padrões de movimentos prováveis podem ser simulados a partir de modelos reais (físicos) para cenários de realidade virtual gerenciados por computador através das tecnologias de processamento de imagem (MAXCY; DRAYER, 2014).

Desde o início dos anos 70, a análise computacional com base científica tem desempenhado um papel vital para o desenvolvimento do registro e análise de dados dos jogos esportivos. Existem estudos científicos que ajudaram a melhorar a compreensão do jogo através da análise de desempenho com resultados teóricos (MEMMERT; PERL, 2009).

O impulso das ferramentas analíticas nos esportes tem origem nos avanços científicos e, portanto, na disponibilidade de grandes quantidades de dados para as equipes e o público em geral. Atualmente, as organizações esportivas são capazes de descobrir, identificar e melhorar o desempenho dos atletas aplicando a análise de dados. O conhecimento é adquirido combinando ciência da computação, estatística e matemática para uma coleção de grandes conjuntos de dados complexos. Equipes como *Oakland A's, Tampa Bay Rays* e *San Antonio Spurs* adotaram o uso da análise. Vale ressaltar que, embora os três clubes sejam considerados pequenos e com recursos limitados, eles obtiveram um enorme sucesso, em parte devido à informação obtida (ALAMAR, 2013).

Os resultados são posteriormente usados para fornecer recomendações e dados relevantes que permitam a otimização ou melhoria no processo de tomada de decisão no esporte, fornecendo, desta forma um número crescente de informações sobre cada atleta. Em esportes altamente competitivos, como o futebol, o nível de desempenho é determinado por um conjunto complexo de variáveis descritas por Weineck (1997), como: (a) a técnica, (b) a tática, (c) fatores psicológicos e (d) a aptidão física.

No entanto, outros tipos de informações, como histórico do atleta, estilo de vida, necessidades psicológicas, podem não ser tão óbvias e evidentes, mas essenciais na optimização do desempenho esportivo do atleta. Alamar (2013), descrevem este conjunto de informações como: (1) Antecedentes, (2) Treinador (3) Testes Psicológicos, (4) Relatórios Médicos, (5) Métricas de Desempenho (6) *Scouting*, (7) Financeiro e (8) Necessidades da equipe.

Indicadores de desempenho específicos podem fornecer informações valiosas, o que pode melhorar a qualidade das intervenções feitas pelos treinadores assim como a otimização do desempenho esportivo de cada atleta. Estes indicadores consistem na seleção ou combinações de diferentes variáveis que visam definir alguns aspectos individuais ou coletivos (HUGHES; BARTLETT, 2002).

A análise do desempenho esportivo constitui a pesquisa sobre o desempenho esportivo real e não as atividades realizadas no laboratório ou através da coleta de dados, tais como questionários, grupos focais, entrevistas, etc. Uma vez que a complexidade do assunto é alta e o esporte é caracterizada pela sua natureza dinâmica, a observação minuciosa e a medição diligente são necessárias para uma melhorar a compreensão desta temática (O'DONOGHUE, 2010).

O principal motivo para o uso da análise de desempenho é superar as limitações da observação subjetiva fornecendo simultaneamente informações objetivas que permitirão uma maior compreensão deste fenômeno. Estas informações irão, por sua vez, auxiliar no processo de tomada de decisão realizado pelos treinadores e comissão técnica. Todo o processo de treinamento é aprimorado e reforçado pela análise de desempenho, segundo Hughes e Franks, (2007).

Uma incrível mudança foi observada em relação à análise de desempenho das equipes, tudo isto graças ao conjunto de ferramentas analíticas utilizadas pelos treinadores e cientistas esportivos (CASTELLANO; ALVAREZ-PASTOR; BRADLEY, 2014; JAMES, 2006; LIEBERMANN et al., 2002). Este desenvolvimento só foi possível em virtude do uso de

métodos analíticos modernos que registram e quantificam os eventos durante o jogo e, posteriormente, geram matrizes de dados que podem traçar as relações de causa-efeito.

Em resumo, é um procedimento normal na maioria dos esportes de elite, obter informações detalhadas do jogo em tempo real, integrando o tempo de movimento e os dados de posição, permitindo assim caracterizar os perfis dos atletas, e com os dados adicionais sobre padrões interpessoais de cooperação que no final suportam as decisões táticas (CARLING; WILLIAMS; REILLY, 2005; TRAVASSOS et al., 2013).

No entanto e apesar de alguns avanços na área de análise de desempenho, a aplicação prática dos resultados da pesquisa infelizmente, está longe da perfeição (CARLING et al., 2013; GARGANTA, 2009; TRAVASSOS et al., 2010; WRIGHT; CARLING; COLLINS, 2014).

A transferência de resultados de pesquisa obtidos no campo para a prática ainda é relativamente rara. Segundo Bishop (2008), a maioria dos estudos não analisa as questões mais relevantes para os profissionais, como o exemplo acima. O estudo de Wright, Atkins e Jones, (2012), afirma que apenas 2% dos treinadores selecionam os principais indicadores de desempenho com base na literatura acadêmica. Assim, revela-se que há uma discrepância significativa entre o foco dos pesquisadores e a aplicação prática realizada pelos treinadores (CARLING et al., 2013 e WRIGHT; CARLING; COLLINS, 2014).

Sem dúvida, existe uma necessidade de melhorar a compreensão do jogo e os treinadores confirmam que o uso da análise de jogos é extremamente importante para melhorar o desempenho de suas próprias equipes, bem como para explorar o adversário. Assim sendo, curiosamente, o foco principal dos treinadores está na análise de vídeo (WRIGHT; ATKINS; JONES, 2012), em que o principal objetivo é melhorar a compreensão do jogo, particularmente para encontrar novas formas de procedimento e novos processos de treinamento.

A análise no futebol consiste na investigação e modelagem do desempenho esportivo, implementando técnicas científicas, mais especificamente o gerenciamento de dados históricos estruturados, a aplicação de modelos analíticos preditivos e a utilização de sistemas de informação, a fim de inteirar os tomadores de decisão e assim permitir vantagem competitiva em campo (MILLER, 2016).

No entanto, ainda existem diversas lacunas no que diz respeito às informações relacionadas ao futebol feminino, tais como características físicas, características de jogos, monitoramento, alterações hormonais, entre outras. Desta forma, esta pesquisa tem como objetivo suprir algumas destas lacunas, tendo como foco principalmente a compreensão e

identificação das principais variáveis individuais, a quantificação das capacidades físicas, assim como a demanda física das atletas de futebol feminino, que de certa forma podem influenciar o seu desempenho esportivo.

DESCRIÇÃO DA SITUAÇÃO DE PROBLEMA

Uma formulação provisória da questão de pesquisa resume-se à formulação de hipóteses, cujas são como "quais são as características de ...?" exigir este elemento (APPOLINÁRIO, 2012).

De acordo com Gil (2008), a análise exploratória dispensa a formulação de hipóteses, mas isso pode ocorrer ao longo do desenvolvimento da pesquisa.

A descrição do problema de pesquisa apresentado neste trabalho se resume à elaboração de uma questão de pesquisa que o pesquisador quer ver respondida na conclusão de sua pesquisa.

Portanto, questões tais como:

I) Quais são as variáveis que devem ser consideradas um time de futebol profissional para determinar o controle das cargas?
II) Quantas variáveis são necessárias para fazer análises de desempenho de jogadores de futebol?

precisam de ser respondidas.

Em outras palavras, é preciso comprentender melhor quais variáveis que definem uma atleta de futebol profissional e mais especificamente que conjunto de variáveis que são importantes no contexto esportivo em questão.

Portanto, este estudo constitui uma contribuição original para o campo da Educação Física, tanto do ponto de vista teórico quanto metodológico. Ele levanta uma discussão sobre métodos avançados de análise do esporte, particularmente a análise realizada no futebol. Além disso, permite criar diretrizes específicas para a análise matemática efetiva de indicadores individuais de atletas.

Além disso, a natureza multidisciplinar do estudo deve ser apontada, pois abrange temas pertencentes às áreas de Gestão da Informação, Educação Física, Matemática, Estatística e Ciência da Informação.

OBJETIVOS DA DISSERTAÇÃO

Os objetivos estão divididos em duas categorias, gerais e específicas, conforme descrição a seguir.

OBJETIVO GERAL

O presente projeto de dissertação tem por objetivo compreender o futebol feminino no Brasil e identificar variáveis-chave individuais importantes numa equipe futebol feminino profissional de alto rendimento, utilizando tecnologias vestíveis.

OBJETIVOS ESPECÍFICOS

Os objetivos específicos derivam do objetivo geral e são os seguintes:

a) Ilustrar o uso potencial de conceitos, técnicas e ferramentas de mineração de textos para identificar lacunas na literatura relativamente a estudos no futebol feminino;
b) Verificar a aplicabilidade do uso de diferentes Sistemas de navegação global por satélite (GNSS) para controle de carga de jogadoras de futebol;
c) Determinar como é o comportamento das demandas de externa das jogadoras futebol durantes as partida em competições diferentes.

ORGANIZAÇÃO DO ESTUDO

Esse trabalho consiste no desenvolvimento de estudos apresentados à posteriori sob o formato de artigos científicos. Desta forma, os artigos serão organizados em forma de capítulos dessa dissertação e serão organidados da seguinte forma:

Capítulo 1. The computational literature review of football performance analysis through probabilistic topic modeling.

Capítulo 2. A systematic review of load control in football using a Global Navigation Satellite System (GNSS)

Capítulo 3. GPS technology to control of external demands of elite Brazilian female football players during competitions

REFERENCIAL TEÓRICO

As atividades esportivas são essenciais para o entendimento das demandas físicas e técnicas (AUGHEY; FALLOON, 2010). O papel da análise de desempenho como a ferramenta de investigação no esporte serve para desenvolver uma compreensão para um esporte em questão, melhorar o desempenho e também auxiliar o processo de tomada de decisão (HUGHES; BARTLETT, 2002; O'DONOGHUE, 2010).

O desempenho no futebol é o resultado de muitas interações dinâmicas (BRADLEY et al., 2011). Carling, Williams e Reilly (2005), explica que no contexto do futebol profissional, a análise de desempenho tem várias aplicações, predominantemente preocupadas com a avaliação tática e técnica, análise de movimento, modelagem de banco de dados e treinamento de treinadores e jogadores.

Inicialmente, o foco da análise era sobre o desempenho físico, enquanto outros estudos menos frequentemente se concentravam no desempenho técnico e tático (RAMPININI et al., 2009; RUSSELL; REES; KINGSLEY, 2013) e demais estudos com foco principal no desenvolvimento e utilização de indicadores para mensuração do desempenho (CARLING; REILLY; WILLIAMS, 2009; CARLING; WILLIAMS; REILLY, 2005; HUGHES; BARTLETT, 2002).

Uma das dificuldades enfrentadas pelos pesquisadores de futebol é estabelecer com sucesso a influência relativa de vários componentes no desempenho dos jogadores (desempenho individual) e equipes (desempenho coletivo) (HUGHES; BARTLETT, 2002). Portanto, é necessário tentar identificar os fatores que afetam significativamente o desempenho atlético (GARGANTA, 2001).

O índice ideal de desempenho atlético deve estar presente enquanto a análise da partida é conduzida para prever os comportamentos futuros da atividade, levando ao sucesso da equipe (JONES; MELLALIEU; JAMES, 2004; O'DONOGHUE, 2005).

Útil do ponto de vista estratégico e tático, a análise estatística do desempenho fornece indicadores que se referem as variáveis de ação que definem um aspecto de desempenho (PRZEDNOWEK et al., 2017).

Quando expressos em termos não dimensionais, os indicadores de desempenho podem ser independentes de qualquer outra variável utilizada (HUGHES; BARTLETT, 2002). Segundo diferentes autores como Jones; Mellalieu e James, (2004), Lago-Peñas; Lago-Ballesteros e Rey (2011) e Williams, Reilly e Reilly (2000), apenas um pequeno número de

estudos tenta fornecer indicadores de desempenho por meio da comparação entre as equipes vencedoras e perdedoras.

Melhorias contínuas na tecnologia e no processo do aprendizado de máquina levam ao progresso no campo da inteligência artificial, permite diversos tipos de análise e seu uso no esporte. A decisão sobre qual é o treinamento de atleta é feita exclusivamente por uma análise dos dados coletados durante a partida. Atualmente, existem vários empresas, equipamentos, softwares e técnicas que buscam os indicadores de desempenho de cada atleta. Entre outros, se precisa destacar:

a) **Técnicas de visão computacional:** são sistemas que capturam automaticamente as posições de cada jogador e da bola em três dimensões várias vezes por segundo e com alta precisão. Esses sistemas capturam a maioria dos jogos nas melhores ligas do mundo (D'ORAZIO; LEO, 2010). Esses dados são combinados com a rotulagem humana do jogo de evento (passes, interceptação, chutes, etc.) e vendidos para as equipes e para a mídia, na forma de um serviço de estatísticas de empresas como Opta, Prozone e Amisco, entre outras (GUDMUNDSSON; WOLLE, 2012; RAMPININI et al., 2007).
b) **Sistema de Posicionamento Global (GPS):** foi aplicado às equipes esportivas na tentativa de superar alguns dos problemas de imprecisão dos dados. O GPS é fornecido como serviço pela Catapult, Stat Sports, GPSports, GPS Pro Soccer, Performa, entre outros. Atualmente percebido como crucial para a compreensão do comportamento e desempenho dos atletas (DAWSON et al., 2004a, 2004b).
c) **Scout:** Várias ações técnicas e táticas da partida podem ser estratificadas por meio de uma análise notacional, como o número total de chutes, chutes para o gol do adversário, gols marcados e sofridos, falta sofrida, impedimentos, os passes certos e errados, tempo de posse de bola e etc. Wyscout, Instant, Sportscode, Videobserver (VO), Nascsport, são, entre outras, as empresas que prestam este serviço (SILVA, 2007).

Este grupo de ferramentas possibilita a coleta de dados de forma substancial para que a equipe de cientistas esportivos possa conduzir as análises relevantes e informar o processo de tomada de decisão do treinador no jogo com base nos indicadores de desempenho (ROBERTSON; BACK E BARTLETT, 2016).

Portanto, o número de variáveis que podem ser coletadas dos atletas usando as tecnologias contemporâneas é suficientemente extenso. Portanto, é importante entender as

principais variáveis de desempenho (individuais e coletivas) que contribuem para o sucesso de uma equipe.

Existem dois tipos principais de indicadores de desempenho: (1) indicador de desempenho individual e (2) indicador de desempenho coletivo.

(1) Indicador de desempenho individual: O sistema de futebol pode ser entendido a partir da decomposição em subsistemas isolados, devido ao fato de depender da interação interna e da interação com o contexto (DAVIDS; ARAUJO; SHUTTLEWORTH, 2005). Dentro deste contexto, como descrito por Weineck, (1997), em esportes competitivos, o nível de desempenho é determinado por um conjunto complexo de variáveis que influenciam os indicadores de desempenho individuais em termos de aspectos técnicos e táticos, bem como da aptidão física. Estes indicadores podem ser entendidos como um conjunto de habilidades dos atletas (técnica individual), em cada posição no campo de jogo (aptidão física) e para cada esquema tático (aspecto tático do jogo).

(2) Indicador de desempenho coletivo: Em esportes que exigem altos níveis de desempenho é fundamental otimizar os resultados com objetivo principal de poder maximizar o desempenho da equipe. Ao longo dos anos, vários autores desenvolveram estudos sobre essa perspectiva no campo da psicologia social buscando a estreita relação com o desempenho da equipe. Dessa forma, estudos nas áreas de motivação, coesão e eficácia coletiva podem ser elencados (CARRON et al., 2002; CARRON; BRAY; EYS, 2002; SMITH; HAMMOND; GILLEARD, 2005).

A relação entre coesão e desempenho tem sido extensivamente estudada por várias áreas (CARRON; BRAY; EYS, 2002). A coesão é definida como um processo dinâmico que reflete a tendência do grupo de se unir e permanecer na busca de seus objetivos e/ou de modo a satisfazer as necessidades emocionais dos membros (CARRON; BRAWLEY; WIDMEYER, 1998). Essa definição leva em conta o aspecto da unidade global e das atrações individuais entre os membros da equipe em relação às tarefas (relacionadas ao desempenho), bem como o aspecto social (orientado pelo relacionamento) da equipe (EYS et al., 2015).

Sobre esses construtos, os seguintes aspectos podem ser identificados: (a) se os jogadores percebem que a equipe atende aos seus objetivos pessoais (atração individual para o grupo), (b) se os jogadores são aceitos pelo grupo (integração individual ao grupo) (c) se os membros da equipe são simpáticos entre si e gostam de comunicação em grupo (tarefa social)

e (d) se trabalham juntos para alcançar objetivos comuns (coesão da tarefa) (CARRON et al., 2002; FUSTER-PARRA et al., 2014; WILSON; OSTROM; COX, 2013).

PRESSUPOSTO TEÓRICO

Existe indícios na literatura sobre o avanço dos estudos relacionados ao futebol feminino, porém verifica-se que o número de estudos ainda é insipiente. Portanto, a partir dessas lacunas são apresentadas as produções cientificas com o objetivo de articular as publicações dentro de uma linha comum e crítica de pensamento. Estes estudos estão aprovados pelo comitê de ética dos Hospital Universitário Pedro Ernesto (HUPE) da Universidade do Estado do Rio de Janeiro (UERJ) sob o número 10529119.8.0000.5259.

Desta forma, será apresentada o artigo intitulado **"*The computational literature review of football performance analysis through probabilistic topic modeling",*** que buscou identificar as lacunas da literatura atravês de uma revisão computadorizada utilizando-se de uma abordagem probabilistica para determinar os tópicos mais estudados no período de 2012 a 2019 sobre análise de desempenho no futebol.

Apartir dos tópicos emergentes encontrados neste primeiro estudo buscou-se aprofundar a temática sobre uso de tencologias durante partidas e treinamentos de futebol dentro do context do futebol feminino no Brasil. Nesta linha de pesamento, o segundo estudo denominado de ***"A systematic review of load control in football using a Global Navigation Satellite System (GNSS)"***, traz uma revisão sobre a quantidade de estudos que utilizam a tecnologia GNSS afim de compreender as demandas físicas das jogadoras de futebol.

Através dos achados e questionamentos trazidos por essa segunda pesquisa foi necessário aprofundar e discutir o comportamentos das varíaveis de carga externa de uma equipe de elite do futebol brasileiro em jogos oficiais em duas competições distintas e concomitantes devido ao calendário de futebol brasileiro. Este atigo tem como título ***"GPS technology to control of external demands of elite Brazilian female football players during competitions".***

1 ESTUDO 1 - THE COMPUTATIONAL LITERATURE REVIEW OF FOOTBALL PERFORMANCE ANALYSIS THROUGH PROBABILISTIC TOPIC MODELING

Abstract

This research aims to illustrate the potential use of concepts, techniques, and tools of process mining to improve the systematic review process. In this way, a review performed on two online databases (Scopus and ISI Web of Science) from 2012 to 2019. We identified 9,649 studies that were analyzed by probabilistic topic modeling procedures in a machine learning approach. The Latent Dirichlet Allocation (LDA) method, chosen for modeling that required the stages: 1) data cleansing, 2) data modeling into topics for coherence, and perplexity analysis. All research was conducted by the standards of the Preferred Reporting Items for Systematic Reviews and Meta-Analyzes (PRISMA) but fully computerized. The computational literature review (CLR) is an integral part of a broader literature review process. The results presented met three criteria: (1) literature review for a research area, (2) analysis and classification of journals, and (3) analysis and classification of academic and individual research teams. A contribution of the article is to demonstrate how the publication's network formed in this particular field of research, and the content of the abstracts can be automatically analyzed to provide a set of research topics for quick understanding and application in future projects.

Keywords Football · Performance Analysis · Literature review · Computational literature review · Topic models · LDA

Mathematical Subject Classification 68T50

JEL Classification C89

Autores: Vitor Principe, Juliana Brandão Pinto de Castro, Luiz Marcelo Carvano, Rodrigo Gomes de Souza Vale, Roberto André Pereira Henriques, Victor José de Almeida e Sousa Lobo e Rodolfo de Alkmim Moreira Nunes

Introduction

Over time, methods for conducting systematic reviews have become more rigorous, further prolonging the completion of reviews (Pham et al. 2018), due to finite resources concerning time and effort (Jennex 2015). Between this, the researcher, a doctoral student, or both new in the research area, need to quickly get an overview of the literature associated with which journals have the most significant impact and witch the most recent and frequent topics (Mortenson and Vidgen 2016).

Thus, researchers contribute to knowledge generation based on searches and promote education. For this, the use of text analysis is beneficial, given the significant increase in the number of electronic research materials in this new era (Lee et al. 2014).

Brings scientists to face new challenges and opportunities due to the characteristics related to the volume, variety, speed of data creation (Chen, Zhong, and Yuan 2016).The systematic literature review (SLR) provides reliable means and established methods for carrying out a comprehensive and robust literature review (Felizardo et al. 2011). However, conducting this researches becomes quite costly due to the studies' growth of 8 to 9% each year, as reported by Bornmann and Mutz (2015). Besides, to being more significant than they used to be, bibliometric datasets are becoming more complex (McLevey and McIlroy-Young 2017).

This abundant data requires computational skills to access these abundant bibliometric data. Several programming languages used to make access easier the academic database APIs, such as projects, *pybliometrics*, Python package (Rose and Kitchin 2019), *rscopus*, R package (Muschelli 2018)to access the RESTful APIs that Scopus provides and other projects can found to access other databases like Web of Science, PubMed, Google Academic and more.

Within this context, after bibliometric data acquisition, Text Mining is a well-established practice. It commonly used to extract non-trivial patterns and knowledge from unstructured documents or textual documents written in natural language (Felizardo et al. 2011).

Among the various methods of text mining and grouping, we highlight probabilistic topic modeling (Blei et al., 2003). This method captures two essential aspects: (1) words can have multiple meanings, and (2) interpretations and documents may contain one or more topics (van Altena et al. 2016).

In this way, natural language processing (NLP) is producing visible practical results due to the advancement of machine learning techniques. One of its main applications is the classification of documents, which received significant attention. In general, document classification problems investigated by (1) coding each word or document for a numerical vector, and (2) classifying documents (Shimada, Kotani, and Iyatomi 2016).

In coding, the Latent Dirichlet Allocation (LDA) method is the most popular topic-modeling algorithm. The LDA attributes the topic distribution of a document probability distribution to the word of each topic (Blei et al. 2003).

For the classification, we highlight the logistic regression, the artificial neural networks, the Bayesian structures, and the support vector machine, which widely used. In recent years, sentence vector representations and recurrent neural networks have shown exciting results in several problems of document classification in English (Shimada et al. 2016).

Therefore, this study aims to demonstrate new essential concepts, mainly for the *Stricto Sensu* programs of Physical Education universities, and to illustrate the potential use of concepts, techniques, and tools of process mining to improve the systematic review process known as computational literature review (CLR). The CLR can identify the main terms and

interpretations found in the articles on soccer performance analysis conducted during the last seven years of scientific production.

Method

The purpose of a literature review is often to allow the researcher to map and evaluate the existing intellectual territory in order to specify a research question and develop additional knowledge (Tranfield, Denyer, and Smart, 2003). However, with the increase in the number of journals, the time and effort required to conduct a literature review are increasing, prompting researchers to choose where to allocate the resources to do empirical research instead of extensive literature reviews. Consequently, the quality outcome of literature reviews is declining (Jennex, 2015).

One possibility to solve the problems of literature reviews is to conduct an SLR, which follows a set of transparent and reproducible steps (an algorithm). In this way, Jahangirian et al. (2011) propose the use of automation to assist in the stages of search and screening.

Research Framework and Development of the Computational Literature Review (CLR)

The available CLR was conducted under the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Liberati et al. 2009) thus providing an overview of the literature and the most relevant topics that published in the studies.

The CLR Framework process (Figure 1) begins with the researcher identifying the type of case (literature review, periodical analysis, research management) that he or she wishes to investigate. For the CLR, the search terms are the same as those used in an SLR.

The structure, shown in Table 1, summarizes the processes and steps of how to extract the latent topics from the data of the articles. The data source for the computational review of the articles were the online databases Scopus and ISI Web of Science, searched on December 1, 2019, for relevant articles published between January 1, 2012, and December 1, 2019, using the keywords “Football,” “Soccer,” each associated with the terms “Performance” and “Analysis.”

Table 1.
Search synthesis

Elements	**Complements**
Period of Analysis	7 years (2012 to 2019)
Databases	Scopus and ISI Web of Science
Keywords	Football, Soccer, Performance, Analysis
Search Method	Bibliometrics (Boolean “AND” and “OR”)
Articles identified	11,413
Articles excluded by repetition	1,764
Articles analyzed	9,649

Created by authors

In order to guarantee the articles’ quality, information sources in the present review were limited to scientific journals. This delimitation is justified, because academics and professionals, to acquire and disseminate knowledge, generally consult scientific journals (Ngai, Xiu, and Chau 2009).

In the selection of the articles, an advanced search procedure used where Boolean expressions ("AND" and "OR") allow combinations of keywords (Rowley and Slack 2004). Then, the articles pre-selected in online journals exported in two different formats:

- Research Information Systems Incorporated (.ris) is a standardized format used by many digital libraries, such as IEEE Xplore, Scopus, ACM Portal, Scopemed, ScienceDirect, SpringerLink, Rayyan QCRI as well as leading reference/citation management applications, such as Zotero, Citavi, Mendeley, and EndNote, which can export and import citations in this format.
- Web of Science Bibliographic Reference (.ciw), currently managed by Clarivate Analytics, used by digital library Web of Science and readable in the bibliographic reference management software of the Clarivate Analytics, which is called EndNote.

Thus, for the organization, identification, and exclusion of duplicate articles, Mendeley Desktop was used because it is free software, easily accessible, and easy to use by researchers. Thus, after deleting duplicate articles, a text file was generated and exported to be used for data analysis.

Impact Analysis

When a research article refers to another article, the original article gets a quote. This way, the number of quotes the article receives can evaluate the impact of an article. Thus, it can be created abstracts by an author (showing how many quotations an author received for the published articles), and by the place of publication (how many quotes a journal received) from counting gross citations to articles included in a CLR. However, citation counts are not problem-free. The h-index used to evaluate the impact of a researcher and generally accepted as a useful measure of impact (Hirsch 2015).

Structure Analysis

Social networks are sets of connected objects represented by a graphic. They have an excellent benefit for the dissemination of information through communication among its members. The network consists of nodes and edges, where each node is a network point, and an edge is a line connecting two nodes (Simsek and Kara 2018; Wasserman and Faust 1994).

A social network reflects a social structure that can be represented by individuals or organizations and their relationships. Through this social structure, data, and information exchanged between individuals or organizations can be studied and analyzed at different levels of detail (Horta et al. 2018).

Scientific Social Networks are specific types of social networks that represent the social interactions of researchers that occur in the scientific environment (Horta et al. 2018). They are very popular in the academic community as a way of understanding the structure of the research community and identifying the top researchers in that community. A component of the scientific authorship and co-authorship network is one in which all authors of this component are reachable (Mortenson and Vidgen 2016).

Content Analysis

In any literature review work, the researcher involved has the concern of identifying the "topics" contained in the documents. In many cases, the evaluation of the work carried out based only on the review of abstracts. In pragmatic terms, this evaluation becomes reasonable because of the amount of work. The abstract purpose: "facilitate quick and accurate

identification of the topic of published papers" (Luhn 1958). The CLR uses probabilistic topic modeling to automate this analysis.

Probabilistic topic models are a collection of algorithmic approaches to machine learning adopted in the field of text mining. These models seek to find structural patterns within a collection of text documents, in order to extract semantic information from a set of documents, called corpus. The topic templates produce groupings of words that represent the central themes present in a particular corpus. In this way, these techniques provide an automated way of identifying common subjects within the documents presented (Lee et al. 2014; Blei 2012; Griffiths and Steyvers 2004).

Given a corpus of documents, probabilistic topic models can find a set of recurring themes called topics. The topics are, in fact, probability distributions on the words of documents. The purpose of topic modeling is to automatically discover topics from a collection of documents (La Rosa et al. 2015).

LDA is a probabilistic statistical model used to discover the underlying abstract topics in a series of documents or text data. (Blei et al. 2003). If it assumed that a document is a sequence *w* of words, where $d = (w_1, w_2, \ldots, w_n)$, the generative model for documents can express through the following probability distribution:

$$P(w_i) = \sum_{j=1}^{T} P\left(w_i \middle| z = z_j\right) P(z = z_j) \quad (1)$$

Where $P(w_i)$ is the probability of the word w_i in a given document; $P\left(z = z_j\right)$ is the probability of choosing a topic word z_j for the current document; $P\left(w_i \middle| z = z_j\right)$ is the probability of showing the word w_i on a certain topic z_j and T is the number of topics. The LDA model is represented as a probabilistic graphical model in Figure 1. This model has been applied in various different fields, such as the detection of topics in collections of articles press (Figuerola et al. 2017). The LDA presents three levels for the representation, where the set of documents is called by the letter D, while $\theta^{(d)}$ is the multinominal distribution on the topics of the document D. The set $N_{(d)}$ denominates the set of words w for a specific document D, while z is the topic to which the word w is assigned. Finally, the set T represents the number of topics, where $\varphi^{(z)}$ is the multinominal distribution on the words for the topic z. For the model called LDA, the latent variables θ, φ and z must be estimated together with the distributed Dirichlet hyperparameters α and β (Blei et al., 2003; Griffiths et al., 2005). The hyperparameters α and β should be interpreted as smoothing factors for assignments respectively from topic to document (θ) and from word to topic (φ).

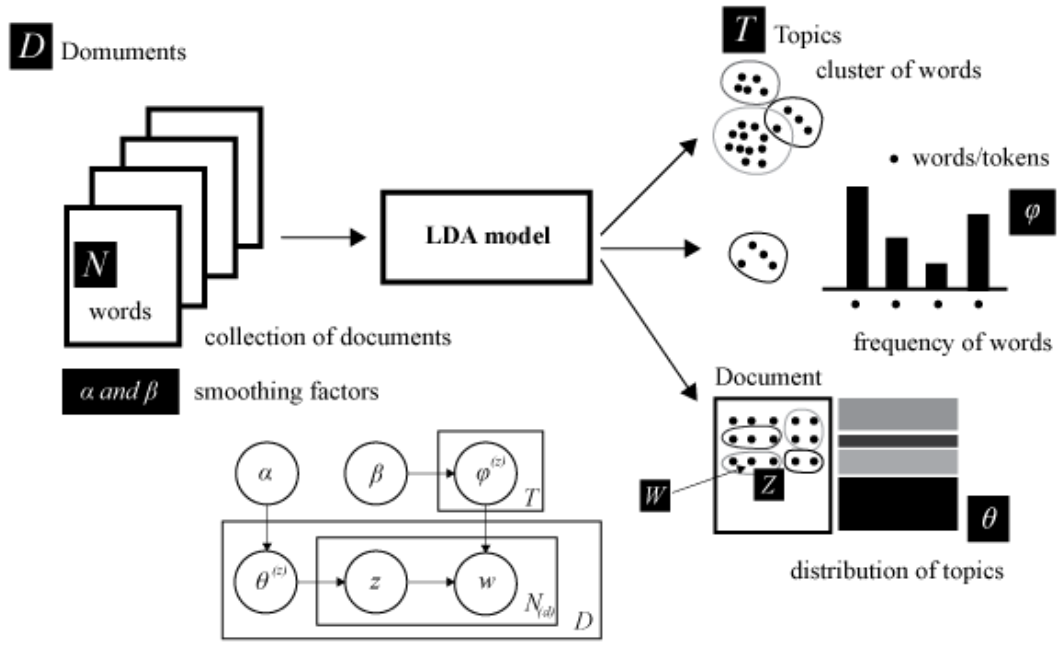


Figure 1. Graphic representation of the LDA model by van Altena et al. (2016). Square boxes represent the sampling as many times as given by the value T, D and N adapted from (Yau et al. 2014; Blei 2012)

Topic Modelling Implementation

The free software Python 3.6 used to implement the steps of pre-processing, topical modeling adjustment, model selection, and post-processing.

Pre-processing of text in this study includes the tokenization of words, conversion of words to upper-case letters, removal of characters and punctuation numbers, and removal of words considered as words of semantic connection (stopwords). Addition, extra stopwords were added, which were garbage words resulting from processing steps.

The assembly of the model and NLP consisted of estimating the latent variables θ, φ, and z, which done using the Gesim 2.2.0 library (Rehurek and Sojka 2010).

Results

The systematic search with time and publication type filters performed using electronic databases (Scopus and ISI Web of Science), last updated December 2019. The search phrase developed with the Boolean operators [OR] (between synonyms) and [AND] (between descriptors). Initially, 11,413 articles identified. After removal of duplicates, 9,649 studies were used (Figure 1).

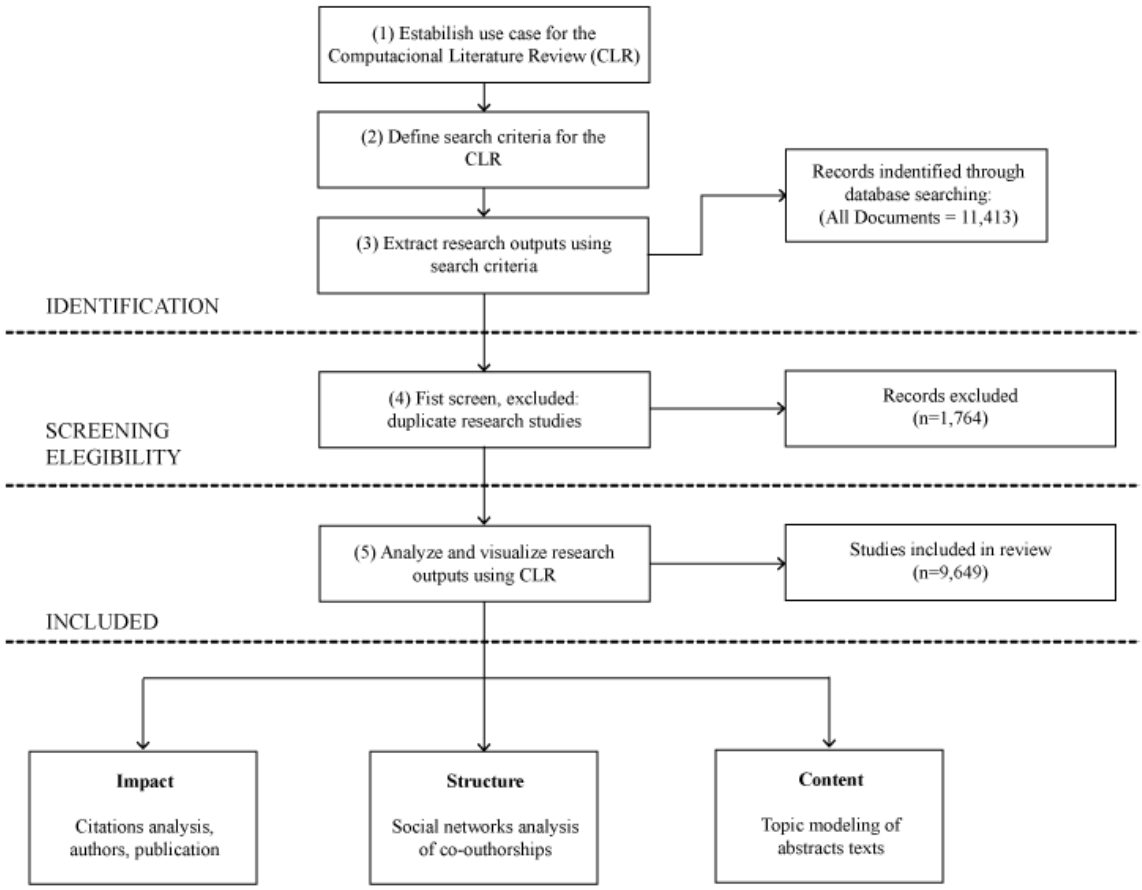


Figure 2. The overview process of the computational literature review adapted from Mortenson and Vidgen (2016). *Notes:* The information flow is described through the different phases of a systematic review according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) model, mapped out the number of records identified, included and excluded, and the reasons for exclusions (Moher et al. 2009)

The inclusion criteria for these articles were: (1) be related to the temporal issue, and thus, the criterion is that the study should be published in the last seven years (2012 to 2019) for analysis, and (2) inclusion of documents solely and exclusively by the type called an article by the two databases. Subsequently, all studies available in the database when searched were selected for this study. Studies also excluded when presented as duplicates.

Impact Analysis

At first, the impact was assessed using the count of publications over the years of articles published in online databases. This is simply the number of articles that were published each year according to Figure 3, and by the journal, as presented in Figure 4.

Table 2 is a pure species of the journals extracted from the databases using the search term(s) sorted by the number of published articles present on that database. This table shows the top 10 journals, although the counts of all journals are written on a spreadsheet so the researcher can conduct further inspections and analyzes.

Then the articles are summarized by the author to identify which researchers have the most significant impact. Table 3 shows the top 10 researchers (out of 9,649 articles) in the data set, according to the number of published works, total citations, and by h-index. Although it is possible to sort data in author order and search for duplicate authors, the

volume of data makes this awkward, and we accept that some "noise" is inevitable. The impact is typically low for author data and has little or no effect on the analysis of location and article citation or in topic modeling of abstracts. It can be seen, in Table 3, that the research in the field of performance analysis in soccer is growing, which shows the interest of several authors on the subject.

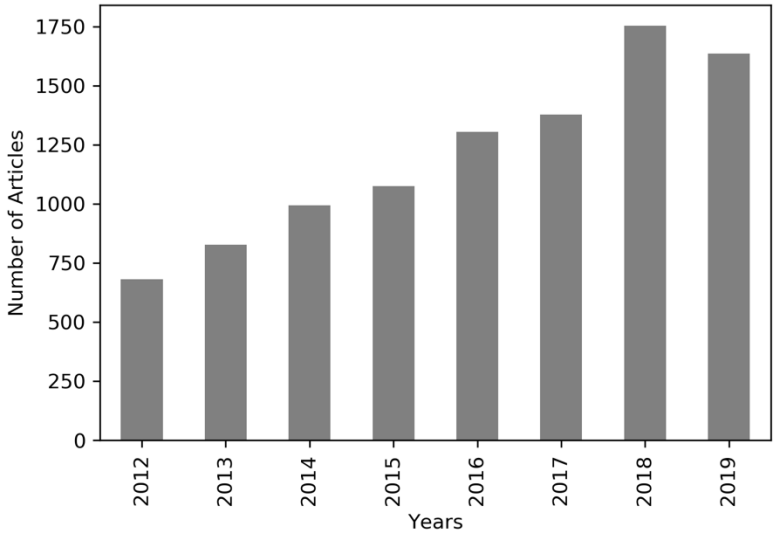


Figure 3. Number of articles publications in the period from 2012 to 2019 in Scopus and ISI Web of Science databases

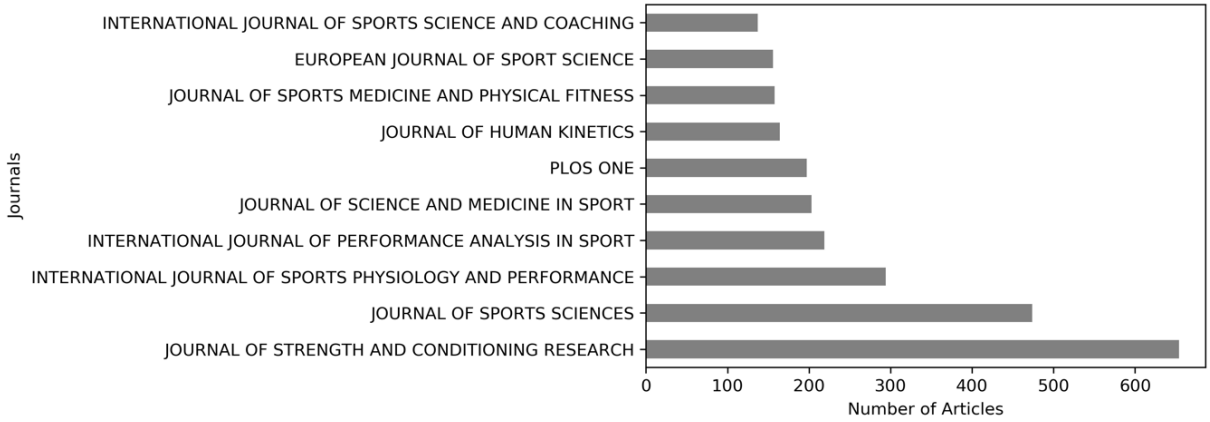


Figure 4. The ten most relevant journals publications for the period 2012 to 2019 based Scopus and ISI Web of Science databases

Figure 4 shows the authors' preference for two databases (Journal of Strength and Conditioning Research and Journal of Sports Sciences) that present more than 470 articles published in this period of analysis and which may indicate a tendency of the themes related to this study area.

In Table 2, the ten journals that obtained the highest number of publications during the period under analysis selected. Therefore, it presented some other impact metrics collected on May 12, 2018, from the respective agencies (Incites Journal Citation Reports, Scimago Journal & Country Rank, and CAPES) in *.cvs* format and included in the database.

Table 2.
Impact of the 10 journals with the most publications in the period from 2012 to 2019

JOURNAL	2019							
	AN	HI	JCR	SJR	CAPES		EF	CITES
					1	2		
Journal of Strength and Conditioning Research	654	108	3.017	1.496	A1	A1	.0176	18,129
Journal of Sports Sciences	474	117	2.811	1.227	A1	A1	.0137	12,711
International Journal of Sports Physiology and Performance	294	40	3.979	1.935	A1	A1	.0096	4,936
International Journal of Performance Analysis in Sport	219	19	1.325	0.646	A2	A4	.0014	1,272
Journal of Science and Medicine in Sport	203	85	3.623	1.665	A1	A1	.0118	6,707
Plos One	197	268	2.776	1.100	A1	A1	1.706	650,727
Journal of Human Kinetics	164	26	1.414	0.644	A2	A4	.0028	1,665
Journal of Sports Medicine and Physical Fitness	158	58	1.302	0.537	A2	B2	.0034	3,247
European Journal of Sport Science	156	41	2.376	1.165	A1	A1	.0066	2,844
International Journal of Sports Science and Coaching	137	19	1.253	0.577	A2	A4	.0013	1,215

AN = articles number, HI = H-index, JCR = Journal Citation Reports, SJR = SCImago Journal Rank, CAPES1 = Old rank by Brazil Higher Education Personnel Improvement Coordination, CAPES2 = New rank by Brazil Higher Education Personnel Improvement Coordination, EF = Eigenfactor Score, Cites = Total number of journal citations

The h-index was created in 2005 by Jorge E. Hirsch as an attempt to measure the impact of academic research. Hirsch (2015) presented an easily computable index, which provides an estimate of the importance, significance, and broad impact of a scientist's contributions, comparing, in an unbiased manner, different individuals competing for the same resource when a critical evaluation criterion is a scientific achievement.

In this way, Plos One is the one that has the most significant impact in the community with an h-index factor of 268 in 2019, many citations, and consequently Eigenfactor Score much higher than the others. Plos One is a free-access scientific journal available only online, published by the Public Library of Science, which mainly covers primary research from any discipline in the field of science and medicine. In this way, Plos One is a journal that needs to be more considered by the authors of this field of study.

Three other journals need to specify attention are Journal of Science and Medicine in Sport, International Journal of Sports Physiology and Performance, and Journal of Strength and Conditioning Research, which have the high impact factors JCR, SJR, and CAPES.

Journal Citation Reports (JCR) is a popular way to evaluate indexed journals on the Web of Science and is a crucial tool to help researchers determine where to publish their work and which journals to use in their research. A little different from the JCR, the SJR indicator is very similar to the Eigenfactor score, the first worked on the Scopus database and the second on the Web of Science database (Jacsó 2010).

By locating the h-index of "someone," several databases can be used. Therefore, for the composition of the data of table 3, the software Harzing's Publish or Perish macOS GU Edition, was used. This software designed to empower academics and present research impact (Harzing 2007). The software can be purchased free of charge from the website https://harzing.com/resources/publish-or-perish. The Publications years established 1990 to 2019 and used search to inspect by Scopus. The Scopus need a free registration required by API Key.

Table 3.
Impact of the ten authors with more publications in the period from 2012 to 2017

AUTHORS	NA	C	HI	INSTITUTION
Clemente, F.M.	706	13,222	50	Polytech Institute of Viana do Castelo, Portugal
Loturco, I.	435	1,040	19	Nucleus of High Performance in Sport, Brazil
Woods, C.T.	354	6,271	34	Port Adelaide Football Club, Australia
Lockie, R.G.	325	1,085	17	California State University, USA
Gabbett, T.J.	263	7,854	52	NR
Delaney, J.A.	238	17,434	64	University of Oregon, USA
Aquino, R.	234	68	5	Federal University of Espirito Santo, Brazil
Castillo, D.	232	98	6	Universidad Isabel I, Spain
Praça, G.M.	225	57	4	Federal University of Minas Gerais, Brazil
Nikolaidis, P.T.	224	1,596	15	University of West Attica, Greece

NA=Number of articles in data base, C=Number of author citation in Scopus, HI=H-index, NR=Not reported

Structure Analysis

In addition to the worksheets used to produce Table 3, the CLR generates a full network view and author views. Figure 5a is a network-wide view of authors and co-authors with 9,650 articles. Figure 5b is a view of the author Clemente, F.M, who presents 700 or more published articles with their respective co-authors in this database. Figures 5c and 5d present in more detail the network of the Clemente, F.M.

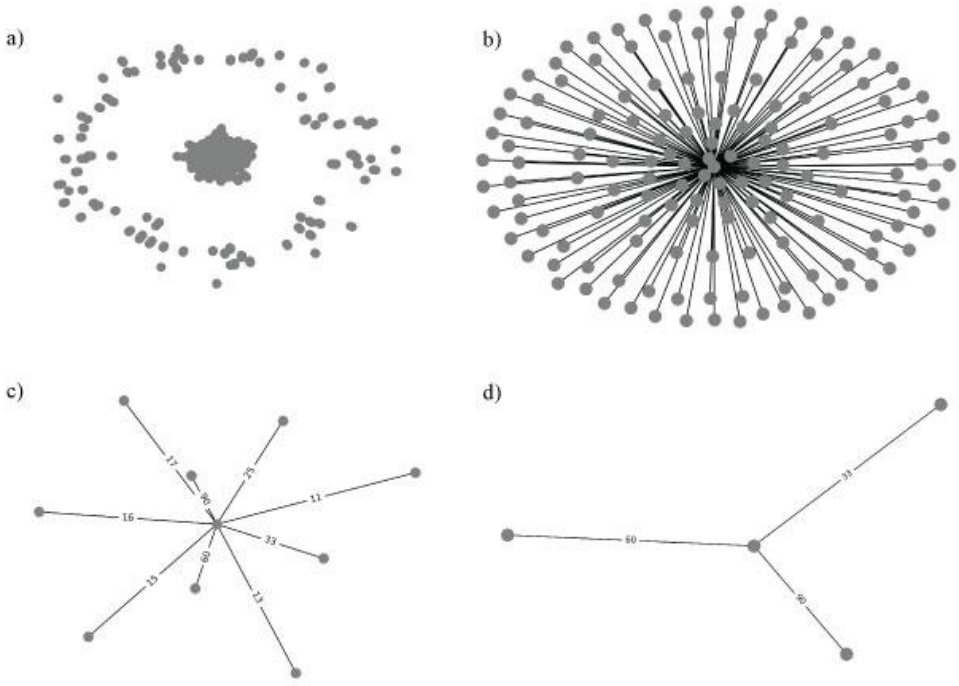


Figure 5. a) Network of all authors in the about football performance analysis period from 2012 to 2019 in Scopus and ISI Web of Science databases b) Network of the Clemente, F.M.

for football performance analysis in Scopus and ISI Web of Science databases c) Network of the author Network of the Clemente, F.M. for football performance analysis in Scopus and ISI Web of Science databases. which has the largest than ten publications d) Network of the biggest contributors (co-author) of the author Clemente, F.M.

It can observe that the authors that research on performance analysis in soccer do not present a homogeneous community, but several segments or niches that are probably determined by their lines of research. It can determine that some authors only develop their work with the same coauthors (the same form of collaboration). However, Clemente, F.M, in his network, presents a more excellent range of publications in partnership, which includes 168 different authors.

Hence, it highlighted Clemente's author, from the Polytech Institute of Viana do Castelo, have secure connections and sharing of works on performance analysis with Martins, F.M.L. (with 88 works together) Mendes, R. (with 60 works published together). Both are Portuguese researchers of the School of Higher Education, of the Polytechnic Institute of Coimbra, Portugal, in the field of Physical Education and Mathematics, respectively, which demonstrates the interest on the application of mathematical models in the analysis of performance in soccer.

Six different nationalities among the ten authors most published from 2012 to 2019, presented in Table 3. Enters Brazil with them the presence of 3 researchers. Highlight Loturco, I. with 1,030 citations based on information collected by Harzing's Publish or Perish software and 435 articles in this database.

Content Analysis

The next task was to build a topic template for abstracts. Required, at first, an extensive data cleansing, which requires relatively little work in this regard, in addition to the standard case, blank, parting, and so on. However, there are still some essential concerns.

The predictive likelihood measures proposed to evaluate the quality of the generated topics. However, its correlation is negative with human interpretation (Chang et al. 2009). In this way, data less consistent with a personal point of view created. This correlation is especially essential when generated topics are used in document collections to understand trends and development within a specific research area (Syed and Spruit 2017). Röder et al. (2015), systematically and empirically explored the topic coherence measures and their correlation with the human topic classification data. Thus, his approach revealed a new measure of unexplored coherence denominate *CV*.

Similarly, Mimno et al. (2011) present a coherence new metric *UMass* where results can classify over the ROC curve area. *UMass* coherence is an asymmetric confirmation measure between major word pairs. Thus, a smoothed conditional probability and perplexity measurement is a predictive measure of the probabilistic model where a low perplexity indicates how good the probability distribution is in the sample (Brown et al. 1992).

First, determine the number of topics (K) to be used. After the analysis of *CV* coherence (0.508), *UMass* (-4.842) and perplexity (-10.566) some experimentation, and considering the size of the data set, we selected a value of 20 (Figure 5).

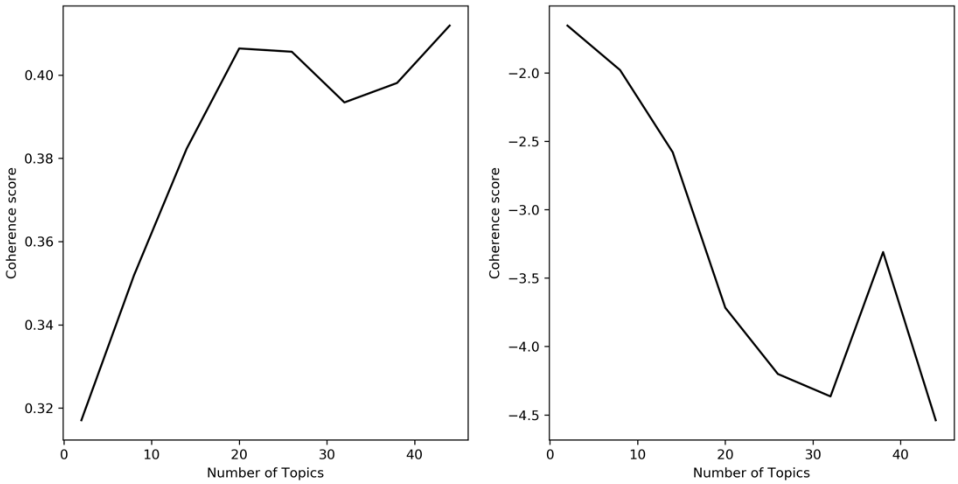


Figure 5. a) *CV* topic model coherence and b) *UMass* topic model coherence

Second, the researcher chooses to remove some words from the data set because they have limited discriminatory value. For example, terms like "performance" or "analysis" occur in almost all extracted abstracts. Additionally, again based on visually inspecting the outputs, we also removed "noise" words, such as "p-value," "American," "et," "role," "however," among others. Although this does not necessarily limit the effectiveness of the model, it makes the results more challenging to interpret, as these terms appear in almost every topic as recommended by Mortenson and Vidgen (2016).

With these preliminary steps performed and defining some of the other required parameters, the model executed. The majority of the 20 topics extracted represent distinct research areas. As an example, it can be seen in Table 4, the first five topics that show the three main most frequent terms for each topic, being the size of the word determined by the probability of the word and human interpretation of the distribution presented by the model.

Table 4.
Top five topics on articles abstract

T	WORD (PROBABILITY)	HUMAN
1	Coaches (~0.009); Development (~0.005); professional (~0.005)	Coaches Support
2	Injury (~0.039); Injuries (0.021); Risk (~0.017)	Injury Risk
3	Physical (~0.0205); Differences (~0.0166); Level (~0.010)	Physical Demands
4	Control (~0.016); Effects (~0.011); Exercise (~0.011)	Exercise Program
5	Information (~0.006); Methods (~0.005); Proposed (~0.005)	Proposed Methods

Presents the three most relevant words with their probabilistic values in the five topics of greatest impact according to the LDA model where T = The topic number and HUMAN = Human interpretation by topics.

Therefore, as in exploratory factor analysis, topic-modeling software does not include label topics – this is something the user must do based on the content of the topics. When working with a large number of documents, we can observe the size of the documents by topic. Thus, Figure 6 shows the number of documents against word distribution. Thus, below (Figure 6) follows the information of 4 topics determined by the research and its relationship with the documents.

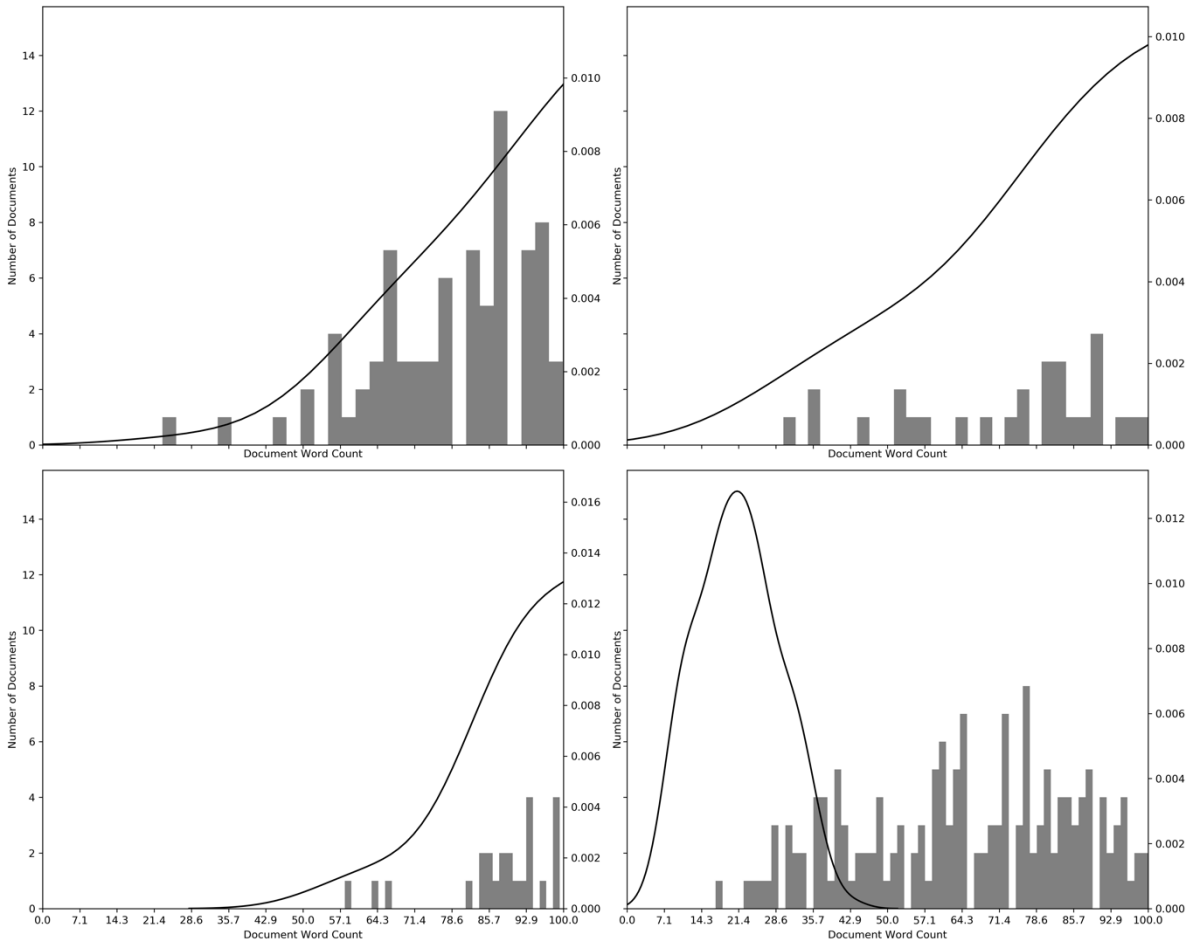


Figure 6. Words distribution in documents publish from 2012 to 2019.

A fascinating visualization technique is provided by the pyLDAvis package, which is a Python library for viewing interactive topic templates based on the package written in R. This package provides an overview of all topics, shows the differences between topics, and allows the researcher to read the most highly associated terms for each topic individually. It is a powerful tool that allows the user to examine specific topics, keeping the entire topic scenario on display, and therefore useful to the user when interpreting and labeling topics (Sievert and Shirley 2014).

Figure 7 shows the pyLDAvis interface. Selecting a topic on the left side (in this case, the topic that seems to address subjects related to physical training types) highlights the most useful terms for interpreting the selected topic on the right side. On the other hand, selecting a term on the right side exposes the conditional distribution on the topics to the left of the selected term.

It is possible to see the research topics related to several areas of performance analysis, such as injuries, strength, and distance, not showing any type of stratification by sex, which demonstrates that the issues through performance analysis in women's football are still incipient. Thus, we highlight topic one related to athlete development and coach support, topic to related to the risk of injury present in all sports, and topic three related to the physical demands on training and games. Other topics can be viewed interactively, for view this use on an IPython notebook, but can also be saved in a standalone HTML file for easy sharing and distribution, as it can check at http://bit.ly/LDA_football

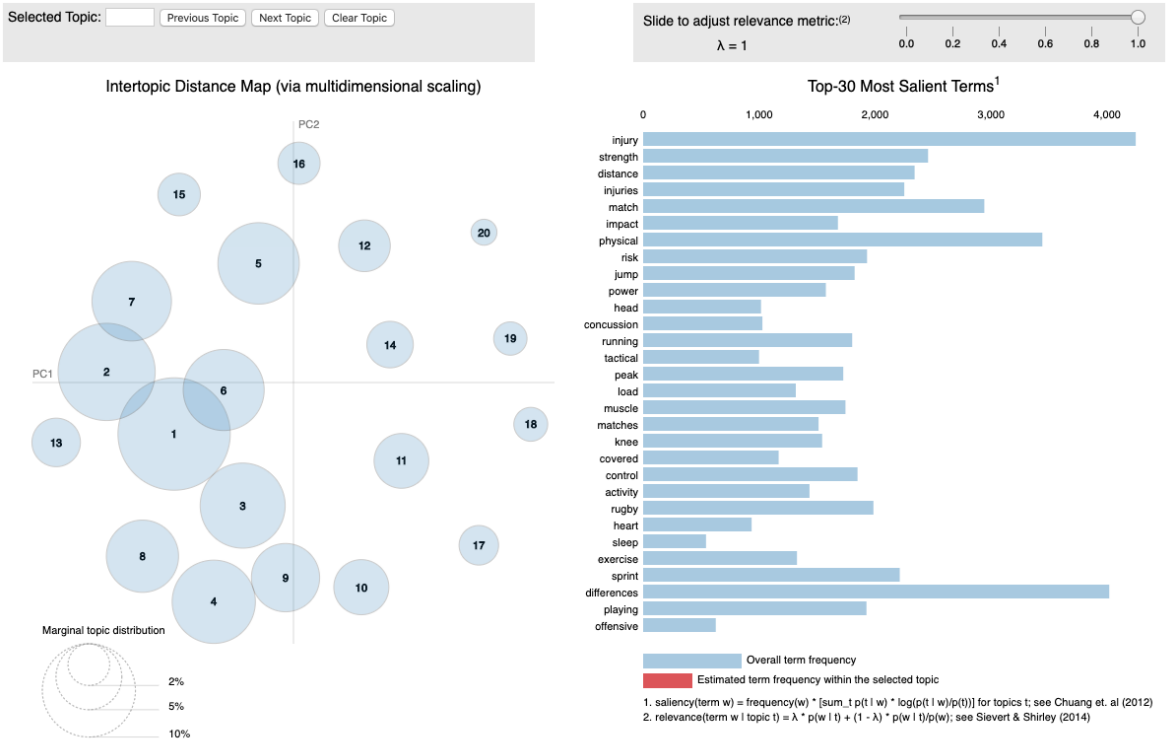


Figure 7. Interactive chart for presentation by pyLDAvis topics (Sievert and Shirley 2014), with a number of publications in the period 2012 to 2019.

After running the topic template, there will be a set of probabilities for each of the articles against the chosen topic numbers. The later probabilities, inferred from the model, demonstrate the "topical distribution" of each document. A document has a probability of 0.5 for a given topic, and this suggests that around 50% of the content of the document is related to that topic.

Study limitations

The study limitations come from a more robust data cleansing model since when working with articles abstracts, the number of 'noise' words is quite high and also the insertion of other sports within the same search context as rugby, Australian football (AFL), handball and others. The missing data and errors in the data insertion in the Scopus and ISI Web of Science database also appear as a limitation.

There is a limitation on the use of only two sources of information (2 online databases) where it does not represent all publications about football performance analysis from 2012 to 2019. Thus, as not using a process more optimized by using the API of these scientific research platform.

Conclusion

As detailed, the CLR offers a fully functional and automated tool that allows the researcher to evaluate large volumes of data on the existing literature regarding impact, structure, and content. Consequently, the CLR offers an approach that may provide greater validity within the academic context on literature reviews. When performed in similar datasets, the CLR results are replicable, and their approach is transparent, providing a more objective way to

determine the relevance and importance of the sources.

Another approach is related to the speed and productivity of research stemming from how academic research can encompass the opportunities of more sophisticated data analysis and the use of large volumes of data in a consistent manner. Although growing in organizations of all kinds, data analysis, conducted using artificial intelligence, has its application in academia, particularly in Physical Education research as an area that is still little explored.

Unusual in the physical education setting, programming languages can assist both systematic review and applicability within the sports context. From this context, the review study identified yet unexplored gaps when considering performance analysis within football. A tiny amount of studies has addressed soccer players, not present for the word woman or significant female probability. Also, we found no studies related to goalkeeper function. Several studies are addressing physical issues, mainly with the use of technologies such as GPS for both professional and young athletes.

As Big Data applications continue to grow in influence in the community as well as the opportunities it offers to conduct new methods of analysis, these professionals' skills may also be more valued. The use of knowledge in the fields of mathematics, machine learning, and artificial intelligence can develop the ability and confidence to use algorithms, through software such as Python, which includes the CLR to support literature reviews with more agility and efficiency.

Acknowledgment

This work was supported by the Carlos Chagas Filho Foundation for Research Support of the State of Rio de Janeiro [grant number E-26/202.638/2018].

References

Blei, D. (2012). Probabilistic topic models. *Communications of the ACM*, *55*(4), 77–84. https://doi.org/10.1109/MSP.2010.938079

Blei, D. M., Edu, B. B., Ng, A. Y., Edu, A. S., Jordan, M. I., & Edu, J. B. (2003). Latent Dirichlet Allocation. *Journal of Machine Learning Research*, *3*, 993–1022. https://doi.org/10.1162/jmlr.2003.3.4-5.993

Bornmann, L., & Mutz, R. (2015). Growth Rates of Modern Science : A Bibliometric Analysis Based on the Number of Publications and cited references. *Journal of the Association for Information Science and Technology*, *66*(11), 2215–2222. https://doi.org/10.1002/asi

Brown, P. E., Pietra, V. J. Della, Mercer, R. L., Pietra, S. a Della, & Lai, J. C. (1992). An Estimate of an Upper Bound for the Entropy of English. *Computational Linguistics*, *10598*(1), 31–40. http://acl.ldc.upenn.edu/J/J92/J92-1002.pdf

Chang, J., Boyd-Graber, J., Gerrish, S., Wang, C., & Blei, D. M. (2009). Reading tea leaves: How humans interpret topic models. *Advances in Neural Information Processing Systems 22 - Proceedings of the 2009 Conference*, 288–296.

Chen, Z., Zhong, F., & Yuan, X. (2016). Framework of Integrated Big Data: A Review. *2016 IEEE International Conference on Big Data Analysis (ICBDA)*, 1–5. https://doi.org/10.1109/ICBDA.2016.7509815

Felizardo, K. R., Salleh, N., Martins, R. M., Mendes, E., MacDonell, S. G., & Maldonado, J.

C. (2011). Using Visual Text Mining to Support the Study Selection Activity in Systematic Literature Reviews. *2011 International Symposium on Empirical Software Engineering and Measurement*, 77–86. https://doi.org/10.1109/ESEM.2011.16

Figuerola, C. G., García Marco, F. J., & Pinto, M. (2017). Mapping the evolution of library and information science (1978–2014) using topic modeling on LISA. *Scientometrics*, *112*(3), 1507–1535. https://doi.org/10.1007/s11192-017-2432-9

Griffiths, T. L., & Steyvers, M. (2004). Finding scientific topics. *Proceedings of the National Academy of Sciences*, *101*(Supplement 1), 5228–5235. https://doi.org/10.1073/pnas.0307752101

Griffiths, T. L., Steyvers, M., Blei, D. M., & Tenenbaum, J. B. (2005). Integrating topics and syntax. *Advances in neural information processing systems*, *17*, 537–544. https://doi.org/10.1.1.73.9813

Harzing, A. W. (2007). Publish or Perish. https://harzing.com/resources/publish-or-perish

Hirsch, J. E. (2015). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(46), 16569–16572. https://doi.org/10.1073/pnas.2336195100

Horta, V., Ströele, V., Braga, R., David, J. M. N., & Campos, F. (2018). Analyzing scientific context of researchers and communities by using complex network and semantic technologies. *Future Generation Computer Systems*, *89*, 584–605. https://doi.org/10.1016/j.future.2018.07.012

Jacsó, P. (2010). Comparison of journal impact rankings in the SCImago Journal & Country Rank and the Journal Citation Reports databases. *Online Information Review*, *34*(4), 642–657. https://doi.org/10.1108/14684521011073034

Jahangirian, M., Eldabi, T., Garg, L., Jun, G. T., Naseer, A., Patel, B., et al. (2011). A rapid review method for extremely large corpora of literature: Applications to the domains of modelling, simulation, and management. *International Journal of Information Management*, *31*(3), 234–243. https://doi.org/10.1016/j.ijinfomgt.2010.07.004

Jennex, M. E. (2015). Literature Reviews and the Review Process : An Editor-in-Chief ' s Perspective. *Communications of the Association for Information Systems*, *36*, 139–146.

La Rosa, M., Fiannaca, A., Rizzo, R., & Urso, A. (2015). Probabilistic topic modeling for the analysis and classification of genomic sequences. *Bmc Bioinformatics*, *16*(Suppl 6), 9. https://doi.org/10.1186/1471-2105-16-s6-s2

Lee, H., Kwak, J., Song, M., & Kim, C. O. (2014). Coherence analysis of research and education using topic modeling. *Scientometrics*, *102*(2), 1119–1137. https://doi.org/10.1021/acsnano.7b00569

Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gøtzsche, P. C., Ioannidis, J. P. A., et al. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: Explanation and elaboration. *PLoS Medicine*, *6*(7). https://doi.org/10.1371/journal.pmed.1000100

Luhn, H. P. (1958). The Automatic Creation of Literature Abstracts. *IBM Journal of Research and Development*, *2*(2), 159–165. https://doi.org/10.1147/rd.22.0159

McLevey, J., & McIlroy-Young, R. (2017). Introducing metaknowledge: Software for computational research in information science, network analysis, and science of science. *Journal of Informetrics*, *11*(1), 176–197. https://doi.org/10.1016/j.joi.2016.12.005

Mimno, D., Wallach, H. M., Talley, E., Leenders, M., & McCallum, A. (2011). Optimizing semantic coherence in topic models. *EMNLP 2011 - Conference on Empirical Methods in Natural Language Processing, Proceedings of the Conference*, (2), 262–272.

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ*, *339*, b2535. https://doi.org/10.1136/bmj.b2535

Mortenson, M. J., & Vidgen, R. (2016). A computational literature review of the technology acceptance model. *International Journal of Information Management*, *36*(6), 1248–1259. https://doi.org/10.1016/j.ijinfomgt.2016.07.007

Muschelli, J. (2018). Gathering bibliometric information from the Scopus API using rscopus. *R Journal*.

Ngai, E. W. T., Xiu, L., & Chau, D. C. K. (2009). Application of data mining techniques in customer relationship management: A literature review and classification. *Expert Systems with Applications*, *36*(2 PART 2), 2592–2602. https://doi.org/10.1016/j.eswa.2008.02.021

Pham, B., Bagheri, E., Rios, P., Pourmasoumi, A., Robson, R. C., Hwee, J., et al. (2018). Improving the conduct of systematic reviews: a process mining perspective. *Journal of Clinical Epidemiology*, *103*, 101–111. https://doi.org/10.1016/j.jclinepi.2018.06.011

Rehurek, R., & Sojka, P. (2010). Software Framework for Topic Modelling with Large Corpora. *Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks*, 45–50. https://doi.org/10.13140/2.1.2393.1847

Röder, M., Both, A., & Hinneburg, A. (2015). Exploring the Space of Topic Coherence Measures. *Proceedings of the Eighth ACM International Conference on Web Search and Data Mining - WSDM '15*, 399–408. https://doi.org/10.1145/2684822.2685324

Rose, M. E., & Kitchin, J. R. (2019). pybliometrics: Scriptable bibliometrics using a Python interface to Scopus. *SoftwareX*, *10*, 100263. https://doi.org/10.1016/j.softx.2019.100263

Rowley, J., & Slack, F. (2004). Conducting a literature review. *Management Research News*, *27*(6), 31–39. https://doi.org/10.1108/01409170410784185

Shimada, D., Kotani, R., & Iyatomi, H. (2016). Document classification through image-based character embedding and wildcard training. *Proceedings - 2016 IEEE International Conference on Big Data, Big Data 2016*, 3922–3927. https://doi.org/10.1109/BigData.2016.7841067

Sievert, C., & Shirley, K. (2014). LDAvis: A method for visualizing and interpreting topics. *Proceedings of the Workshop on Interactive Language Learning, Visualization, and Interfaces*, 63–70. https://doi.org/10.1.1.100.1089

Simsek, A., & Kara, R. (2018). Using swarm intelligence algorithms to detect influential individuals for influence maximization in social networks. *Expert Systems with Applications*, *114*, 224–236. https://doi.org/10.1016/j.eswa.2018.07.038

Syed, S., & Spruit, M. (2017). Full-Text or abstract? Examining topic coherence scores using latent dirichlet allocation. *Proceedings - 2017 International Conference on Data Science and Advanced Analytics, DSAA 2017*, *2018-January*, 165–174. https://doi.org/10.1109/DSAA.2017.61

Tranfield, D., Denyer, D., & Smart, P. (2003). Towards a Methodology for Developing Evidence-Informed Management Knowledge by Means of Systematic Review, *14*, 207–222. https://doi.org/10.1111/1467-8551.00375

van Altena, A. J., Moerland, P. D., Zwinderman, A. H., & Olabarriaga, S. D. (2016). Understanding big data themes from scientific biomedical literature through topic modeling. *Journal of Big Data*, *3*(1). https://doi.org/10.1186/s40537-016-0057-0

Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications*. Cambridge University Press.

Yau, C. K., Porter, A., Newman, N., & Suominen, A. (2014). Clustering scientific documents with topic modeling. *Scientometrics*, *100*(3), 767–786. https://doi.org/10.1007/s11192-014-1321-8

2 ESTUDO 2 - A SYSTEMATIC REVIEW OF LOAD CONTROL IN FOOTBALL USING A GLOBAL NAVIGATION SATELLITE SYSTEM (GNSS)

Abstract

Aim: The purpose of this study is to verify the applicability of using different global navigation satellite systems (GNSS), from frequency applied, the number of satellites available and the dilution of precision (DOP) presented to measure football players load control. **Method:** The systematic review followed the PRISMA recommendations. Four hundred and eighty-five articles were selected from two online databases (Scopus and ISI Web of Science) over five years, with 21 studies selected for this review. In these studies, the GNSS frequency ranged from 5 to 18 Hz, with the most commonly used 10 Hz frequency. **Results:** Out of 21 selected studies, 20 studies of those presented the ideal horizontal dilution of precision (HDOP), and the number of satellites available varied between 5 to 20. There were no lines for defining speed, acceleration, and deceleration zones. **Conclusion:** There was no uniformity in data collected from devices. Most GNSS devices brands do not adopt the international system of units (SI).

Palavras-Chave: Load Control, Football, Global Navigation Satellite System, Systematic Review, GNSS, GPS

Introduction

The load control in training and game sessions is an essential strategy for reducing injuries and optimizing performance in football[1]. When quantifying loads of athletes, it is possible to observe the amount of activity performed (athlete external load) as the primary determinant for individual physiological responses (athlete internal load)[2].

Load control standard parameters have become part of the everyday life of football clubs in the world[3], considering that from 2015 the International Football Association Board allowed the use of wearable technologies during competitions or official matches.

Thus, quantifying the external loads in football includes evaluating distances, reached speeds, performed accelerations, and decelerations, as well as the exposure time performed by the athlete[4]. These external loads can be determined with the aid of electronic tracking systems that are differentiating into 1) external positioning system (OPS), 2) local positioning systems (LPS), and 3) image-based systems[5].

Through OPS, the Global Navigation Satellite System (GNSS) calculates the athlete's position through trigonometry between satellites. The accuracy of the information depends on factors such as the number of satellites connected to each device and the dilution of precision (DOP), which provide a measurement accuracy of the GNSS signal, determined by the position of satellites in the sky[6].

The satellites, which are far away in the sky, provide the lowest error triangulation. Consequently, one PDO lowers, while the closest satellites result in a significant error triangulation, with a higher DOP. If more than four satellites are used, a three-dimensional location is calculated[7].

GNSS systems establish autonomous geospatial positioning through the use of artificial satellites. In this sense, the US global positioning system (GPS) and Global Navigation Satellite System by Russian are considered fully operational and global reach (GLONASS)[8].

In addition to the technological development and the enormous amount of research using GNSS devices[9,10], the extensive use of this technology deserves further scientific deepening for analysis validity and reliability of the system used.

Thus, this systematic review aimed to evaluate the applicability of the use of various Global Navigation Satellite Systems (GNSS), known in the market as GPS from the applied sampling frequency, number of satellites available and the dilution of precision (DOP) presented to measure load control in football athletes.

Methods

A systematic review of the literature from two online databases available on the load control of football players was held using a Global Navigation Satellite System (GNSS) to measure the training loads or games. The study are approved by the Research Ethics Committee of Hospital Universitário Pedro Ernesto (HUPE) of the Rio de Janeiro State University of (UERJ) under number 10529119.8.0000.5259.

Study Design

This systematic review was conducted according to the guidelines of PRISMA (Preferred Reporting Items for Systematic Review and Meta-analysis)[11]. The searches were conducted in two electronic databases (Scopus and ISI Web of Science) to identify articles between January 2015 and December 2019. The related search terms were: Soccer, Football,

Load Control Monitoring Load, Player load, Training Load, Matches Load, Global Position System, GPS, and Wearable Technology. Boolean operators 'OR' and 'AND' were used to construct the final search phrase (Table 1).

Table 1. Search synthesis

Elements	Complements
Period of Analysis	5 years (2015 to 2019)
Databases	Scopus and ISI Web of Science
Keywords	Football, Soccer, Load Control, Monitoring Load, Player Load, Training Load, Matches Load, Global Position System, GPS, Wearable Technology
Search Method	Bibliometrics (Boolean “AND” and “OR”)
Articles identified	385
Articles excluded by repetition	74
Articles out of scope	390
Articles analyzed	21

Study Eligibility and Inclusion Criteria

All results in an electronic search are initially exported to Mendeley® Desktop software. The identification of eligible studies followed a three-step process: (1) duplicate studies were removed, (2) studies that were out of scope were excluded after screening title and abstract if an assertive decision could not be made in this phase, the studies were carried out and (3) the last stage was completed by three independent researchers, which involved the removal of the studies by the exclusion criteria after the complete screening of the text. The present study adopted the following inclusion criteria: 1) The study written in English; 2) The study published as an original research in a peer-reviewed journal; 3) Data reported only for football; 4) The sample consisting of competitive athletes (defined as by university students, college, international, professional, experienced, semi-professional, national athletes); and 5) The use of a global navigation satellite system (GNSS) is mandatory.

Bias Analysis

The quality of the studies was evaluated through the risk of bias using the procedures adopted by[12], as shown in Table 2. The scores were assigned based on how much each criterion has been met, until a maximum possible score of 8 (low risk of bias). Studies with a risk of bias score equal to or less than four were considered inadequate and were excluded. Once selected the studies to be included, there was a review of the check reference lists[13] to identify additional peer-reviewed studies.

Table 2. Risk of bias

	Criteria	Definition	Scoring		
			0	1	2
A	Peer-reviewed	Study publish in peer-reviewed journal	No	Yes	
B	Number of participants	Number of players include in study findings	<5	5-50	>50
C	Population defined	Age, gender, sport, participation level and experience stated	No	Partly	Yes
D	Training or competition load described	Training or competition undertaken during the study period	No	Partly	Yes
E	GPS use	The use of GPS was described	No	Yes	

Adapted from Horsley et al. (2009)[13]

Data Extraction

After meeting the eligibility criteria for the inclusion of the studies, the data of the selected articles were organized in a personalized spreadsheet of Microsoft® Excel version 16.31 for Mac. However, according to recent recommendations from[9,14], only articles that presented the records and details of the brand and = device model, data acquisition frequency (Hz), number of satellites, horizontal dilution of precision (HDOP) and software used for data acquisition, were analyzed and included in this review.

Table 3. Effect of dilution of precision

HDOP Level	Quality / Signals		Description
< 1	Ideal	⇑⇑	Maximum accuracy possible all times
1-2	Excellent	⇑	Accurate measurements
2-5	Good	⇒	Measurements with appropriate precision
5-10	Moderate	⇔	Moderate quality, recommended to fix
10-20	Weak	⇓	Consider discarding all data
>20	Poor	⇓⇓	Very low accuracy, errors can around 300m

Adapted from Moen (1997)[7] and Wu et al (2012)[40].

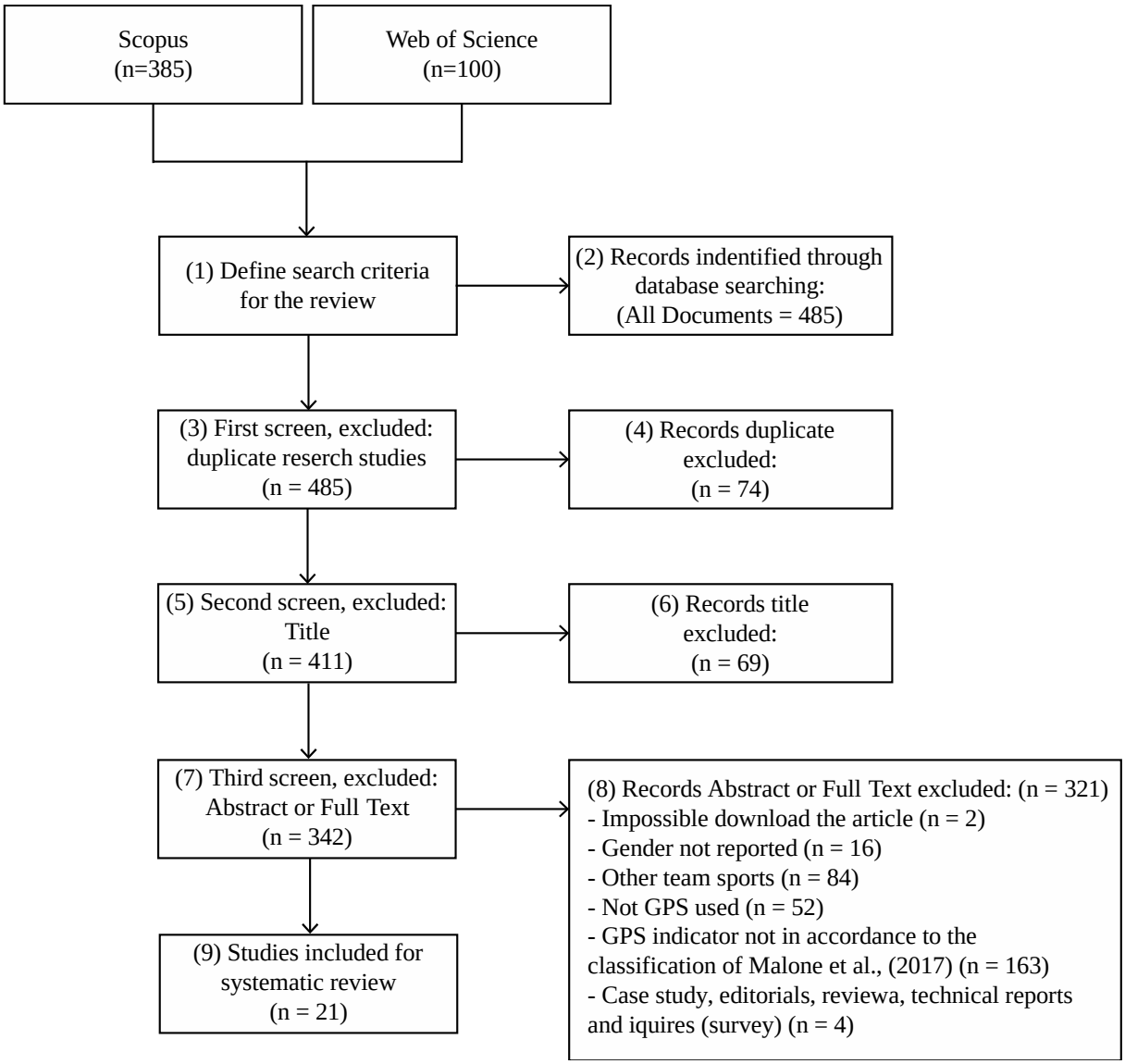
Results

Studies Overview

A systematic search was performed using time filters, publication type, and language using electronic databases Scopus and ISI Web of Science, with the latest update in December 2019. Descriptors in Health Sciences were used (DeCS) and Medical Subject Headings (MeSH). The search phrase was developed with the Boolean operators [OR] (between synonyms) and [AND] (between descriptors).

Initially, 485 records were identified. After removing duplicates, download impossibilities and screening the titles and abstracts, 320 studies met the inclusion criteria. During the screening of the full text, the main reasons for exclusion were: other sports different from football, genre uninformed, other different technologies that used GPS, unoriginal studies (case study, editorial, review or technical reports), resulting in a total of 21 studies included (Figure 1).

Figure 1. The systematic review information flow described by phases according to the preferred reporting items for systematic reviews and meta-analyzes (PRISMA)



Description and quality of the studies

The research had been published from 2015 to 2019 in different leagues with different ages. Out of 21 studies, 14 were performed with male football players, five studies with female football players, and one study with both genders, as shown in Figure 2. Among the 21 selected studies, the average score on the risk of bias was 6.9, in which no article reached the maximum score of 8 points.

The deficiency presented was the non-inclusion of the age of the 45 athletes involved in the study by Trewin et al. (2018)[15], in which researchers only reported that the athletes were part of a senior national team ranked in the top 10 of the world and the duration of some studies. Madison et al. (2019)[16] portray their sample as semi-professional football players, and Scott and Lovell (2018)[2] presented an interim team for the FIFA women's world cup, which can be seen in greater detail in Table 4.

Among the 21 studies selected, data from the GPS system were collected in 1 day of training up to two full seasons, where 75% of the articles carried out their investigations based on training, with some studies during the competitive period of the athletes and 25% using of official games for data collection, according to Table 5. None of these studies included data from goalkeepers. Research by Knight et al. (2015)[17] and Trewin et al. (2018)[15] did not limit the time taken to acquire GPS data for their studies.

It is highlighted in the selected studies the control of external load mainly in training on the physical conditioning of the athletes (~ 38%), either in the control of the activities proposed by the physical trainer or by physical tests applied for the specific use of the presented study. Game applications are present in ~ 47% of studies presented as competitive games or official games (Table 5).

Figure 2. Research frequency of football load control

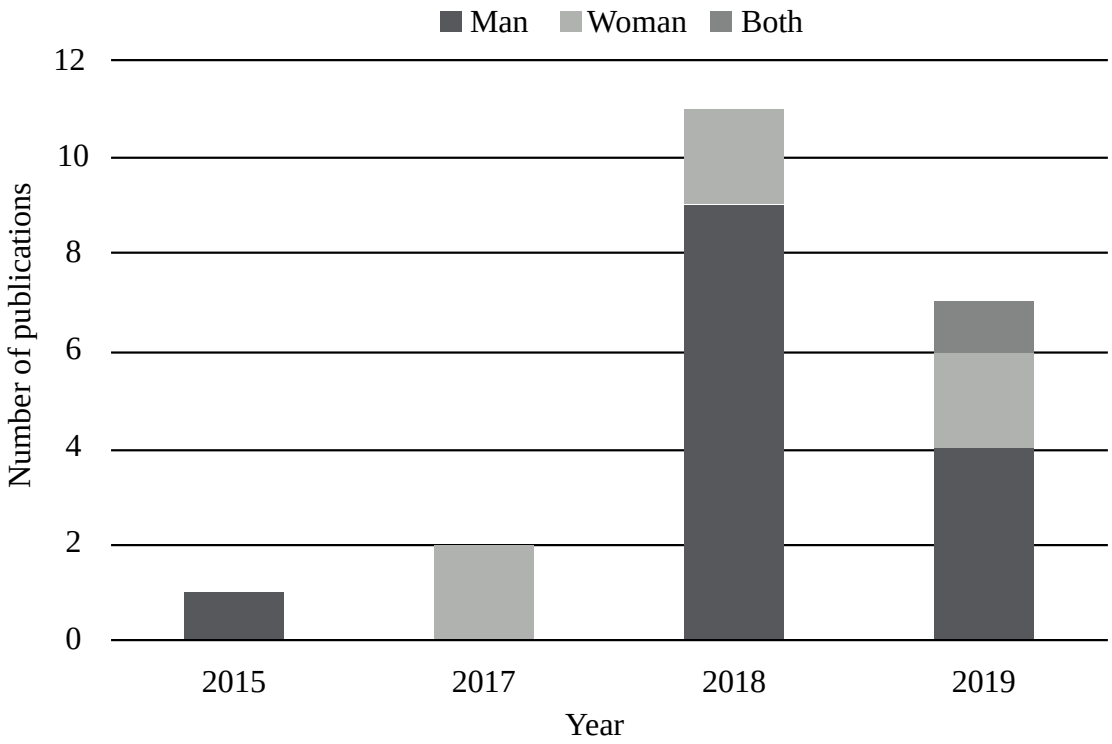


Table 4. Quality Description of the 21 selected studies

Authors	Gender	Number of Players (Age)	Context	Risk of Bias					
				A	B	C	D	E	Total
Abbott et al. (2018)[22]	M	46 (19.1 ± 1.2)	U23 Premier League Academy	1	1	2	2	1	7
Abbott et al. (2018)[23]	M	31 (19.4 ± 1.7)	U23 Premier League Academy	1	1	2	2	1	7
Abbott et al. (2018)[24]	M	19 (18.2 ± 1.1)	U21 Premier League Academy	1	1	2	2	1	7
Abbott et al. (2018)[25]	M	37 (19.9 ± 1.4)	U23 Premier League Academy	1	1	2	2	1	7
Beato et al. (2019)[29]	M/F	20 (21.0 ± 2.0)	University Students	1	1	2	2	1	7
Beato et al. (2018)[40]	M	15 (20.0 ± 1.0)	University Students	1	1	2	2	1	7
Curtis et al. (2018)[30]	M	18 (20 ± 1)	NCAA Division I	1	1	2	2	1	7
Fitzpatrick et al. (2018)[29]	M	14 (17.1 ± 0.5)	U18 Premier League	1	1	2	2	1	7
Jaspers et al. (2018)[31]	M	35 (23.2 ± 3.7)	Netherlands First league	1	1	2	2	1	7
Jones et al. (2019)[34]	M	37 (23 ± 4)	English Football League One	1	1	2	2	1	7
Knight et al. (2015)[17]	M	14 (22 ± 4.1)	Australian domestic	1	1	2	2	1	7
Kyprianou et al. (2019)[41]	M	12 (16.3 ± 0.8)	Elite youth football academy	1	1	2	2	1	7
Madison et al. (2018)[16]	M	10 (23 ± 5)	NR[1]	1	1	2	2	1	7
Massard et al. (2018)[42]	M	23 (21.4 ± 3.8)	Australian Tier 2 club	1	1	2	2	1	7
Meylan et al. (2017)[33]	F	20 (18.2 ± 0.7)	U20 women's national team	1	1	2	2	1	7
Park et al. (2019)[38]	F	27(24.6 ± 3.8)	USA women team	1	1	2	2	1	7
Rago et al. (2019)[32]	M	13 (25.8 ± 3.5) 17 (15.6 ± 0.5)	Italy's second-tier league	1	1	2	2	1	7
Ramos et al. (2019)[28]	F	14 (18.1 ± 0.8) 17 (27 ± 4.5)	Brazilian national teams	1	1	2	2	1	7
Ramos et al. (2017)[27]	F	12 (18 ± 0.7)	U20 women's Brazilian team	1	1	2	2	1	7
Scott et al. (2018)[2]	F	22 (21.9 to 39.5)	NR[2]	1	1	2	2	1	7
Trewin et al. (2018)[15]	F	45 (NR)	NR[3]	1	1	1	1	1	5
			All studies mean	1	1	1.95	1.95	1	6.9

M=Male, F=female, NR =Not reported, NR[1]=Semiprofessional football players, NR[2]=Representing a provisional squad for the 2015 Women's FIFA World Cup, NR[3]=senior national team ranked top 10 in the world, A: Study publish in peer-reviewed journal (No=0, Yes=1), B: Number of players include in study findings(<5=0, 5-50=1, >50=1),, C=Age, gender, sport, participation level and experience stated (No=0, Partly=1, Yes=1), D: Training or competition undertaken during the study period (No=0, Partly=1, Yes=1), E: The use of GPS was described (No=0, Yes=1).

Table 5. GPS data duration and application

Authors	Duration	Training Session						Matches		
		PC	TeP	TP	SSG	D	RC	F	C	O
Abbott et al. (2018)[22]	August 2016 to May 2017					YES				
Abbott et al. (2018)[23]	4 weeks in pre-season period	YES	YES	YES	YES		YES	YES		
Abbott et al. (2018)[24]	6 weeks in pre-season period	YES						YES		
Abbott et al. (2018)[25]	6 weeks in pre-season period	YES	YES					YES		
Beato et al. (2019)[29]	Data recorded in 2018	YES								
Beato et al. (2018)[40]	NR	YES				YES				
Curtis et al. (2018)[30]	2015 (August to November)									YES
Fitzpatrick et al. (2018)[29]	6 weeks (August-September)								YES	
Jaspers et al. (2018)[31]	2 seasons 2014–2015 and 2015–2016								YES	
Jones et al. (2019)[34]	2 seasons 2015-2016 and 2016-2017								YES	
Knight et al. (2015)[17]	NR				YES	YES				
Kyprianou et al. (2019)[41]	8 training sessions per week								YES	
Madison et al. (2018)[16]	4 weeks				YES					
Massard et al. (2018)[42]	Pre-season	YES								YES
Meylan et al. (2017)[33]	34 games							YES		
Park et al. (2019)[38]	Games between 2012 and 2015									YES
Rago et al. (2019)[32]	15–25 training sessions	YES	YES	YES				YES		
Ramos et al. (2019)[28]	Official international competitions									YES
Ramos et al. (2017)[27]	U-20 South American Championship									YES
Scott et al. (2018)[2]	1-day training camp	YES								
Trewin et al. (2018)[15]	NR									YES

NR=Not reported, PC=Physical Conditioning, TeP=Technical Practices TP=Tactical Practices, SSG=Small-side Games, D=Drills, RC=Replication of Competition, F=Friendly Matches, C=Competitive Matches, O=Official Matches.

GPS Variables

For ~ 66% of the studies, the total distance adopted represented the control of the training volume. Out of these studies, ~ 33% also used the total number of accelerations for the same purpose and only ~ 23% used the total number of decelerations. Thus, it is worth noting that the studies diverged on how to quantify the accelerations and decelerations, while 10 articles selected in this study were also concerned with some control of the internal load of the athletes during the execution of the activities and their correlation with the data of external load acquired with the aid of GPS. About ~ 33% of the studies used HR for this control, ~ 14% used the RPE scale, ~ 9% sometimes measured the athletes' oxygen volume, and ~ 4% assessed the lactate concentration in the blood. An essential organization of the variables adopted by the studies is shown in Table 6.

Table 6. External and internal load application

Authors	External Load																												
	Total									Running				Acc.			Dec.			Exp.									
	TD	TA	Td	SD	WD	JD	RD	FRD	HMD	VHI	HI	MI	LI	HI	MI	LI	HI	MI	LI	HI	MI	S	PS	AMP	AS	MA	RV	PL
Abbott et al. (2018)[22]	Y									Y				Y	Y		Y	Y		Y	Y	Y							
Abbott et al. (2018)[23]Φ														Y	X	X										Y			
Abbott et al. (2018)[24]∂										Y	Y											Y							
Abbott et al. (2018)[25]	Y									Y				Y	X							Y							
Beato et al. (2019)[29]Φ	Y																						Y						
Beato et al. (2018)[40]∂Φ	Y								Y															Y			Y		
Curtis et al. (2018)[30]∂Φ	Y				Y	Y	Y				Y			Y	Y	Y	Y	Y	Y			Y			Y				Y
Fitzpatrick et al. (2018)[29]∂	Y	Y	Y							Y	Y																		
Jaspers et al. (2018)[31]∂Φ	Y	Y	Y								Y																		
Jones et al. (2019)[34]	Y											Y	Y									Y							Y
Knight et al. (2015)[17]∂	Y			Y	Y	Y	Y	Y			Y		Y									Y			Y				
Kyprianou et al. (2019)[41]Φ																						Y							
Madison et al. (2018)[16]∂Φ	Y	Y	Y																										
Massard et al. (2018)[42]Φ																						Y							
Meylan et al. (2017)[33]	Y	Y									Y			Y								Y							
Park et al. (2019)[38]										Y	Y		Y									Y	Y						
Rago et al. (2019)[32]∂											Y	Y										Y							
Ramos et al. (2019)[28]Φ	Y	Y	Y								Y											Y							Y
Ramos et al. (2017)[27]∂	Y	Y	Y																			Y							Y
Scott et al. (2018)[2]∂										Y	Y												Y						
Trewin et al. (2018)[15]Φ	Y	Y									Y		Y									Y							Y

∂=Internal load data was collected, Φ=Practical applications was described, Y=Yes, Acc.=Acceleration, Dec.= Deceleration, Exp.=Explosive, TD=Number total distance, TA= Number total acceleration, Td= Number total deaceleration, SD=Standing distance, WD=Walking distance, JD=Jog distance, RD=Run distance, FRD=Fast running, HMD=High metabolic distance, VHI=Very high intensity distance, HI= High intensity distance, MI= Moderate intensity distance, LI= Low intensity distance, S=Sprinting, PS=Peak Speed, AMP=Average Metabolic Power, AS=Average speed, MA=maximum acceleration, RV=Relative velocity, PL=Player Load

GPS Information

Catapult® is the company that supplied most of the equipment for measuring data. About ~ 80% of the articles found in this study used equipment from this company, and its MinimaxX model represents ~ 42% of use, surpassing the most recent Vector models that do not appear in any study and the Optimeye line used in ~ 28% of the studies. The frequency of data acquisition was mostly 10Hz, with the presence of one study using the frequency of 5Hz and two studies using the Apex 18Hz device from STATsports®, as shown in Table 7.

The frequency of data acquisition was mostly 10Hz, with the presence of one study using the frequency of 5Hz and two studies using the Apex 18Hz device from STATsports®. It is noteworthy that the accuracy of GPS devices has been extensively investigated, and has shown an improved precision with a sampling frequency of up to 10Hz[18,19], there are no additional benefits when increasing the frequency acquisition[20].

Table 7. Devices information's

Authors	GPS					
	Company	Model	Frequency	Number of Satellite	HDOP	Software
Abbott et al. (2018)[22]	Catapult	MinimaxX 4.0 / GPS	10Hz	> 12	<1 ⇑⇑ (ideal quality)	Sprint 5.1.5
Abbott et al. (2018)[23]	Catapult	OptimEye S5B / GPS	10Hz	15 ± 1	0.8 ± 1 ⇑⇑ (ideal quality)	Sprint 5.1.5
Abbott et al. (2018)[24]	Catapult	OptimEye S5B / GPS	10Hz	15 ± 1	0.7 ± 1 ⇑⇑ (ideal quality)	Sprint 5.1.5
Abbott et al. (2018)[25]	Catapult	OptimEye S5B / GPS	10Hz	14.7 ± 1.8	0.8 ± 0.1 ⇑⇑ (ideal quality)	Sprint 5.1.5
Beato et al. (2019)[29]	STATsports	Apex / multi-GNSS	10Hz 18Hz	18 ± 2	0.4 ± 0 ⇑⇑ (ideal quality)	Apex 2.0.2.4 Apex 5.0
Beato et al. (2018)[40]	STATsports	Apex / multi-GNSS	10Hz	18 ± 2	0.4 ± 0 ⇑⇑ (ideal quality)	Apex 2.0.2.4
Curtis et al. (2018)[30]	Catapult	MinimaxX 4.0 / GPS	10Hz	9 ± 3	1.6 ± 1.3 ⇑⇑ (ideal quality)	Logan Plus 5.1.7
Fitzpatrick et al. (2018)[29]	Catapult	MinimaxX S4 / GPS	10Hz	14.4 ± 0.5	0.81 ± 0.10 ⇑⇑ (ideal quality)	Sprint 5.1.7
Jaspers et al. (2018)[31]	Catapult	Minimax S4 / GPS Optimeye S5 / GPS	10Hz	≥ 8	< 1.5 ⇑ (excellent quality)	Sprint 5.1.7
Jones et al. (2019)[34]	Catapult	Optimeye X4 /GPS	10Hz	≥ 6	< 1.5 ⇑ (excellent quality)	Sprint 5.1.7 Openfiel 1.11.2
Knight et al. (2015)[17]	Catapult	MinimaxX/ GPS	5Hz	> 5	< 8 ⇑⇑ (ideal quality)	Logan Plus 4.4.0
Kyprianou et al. (2019)[41]	Catapult	Optimeye S5/ GPS	10Hz	12.0 ± 0.0	0.68 ± 0.04 ⇑⇑ (ideal quality)	Openfield 1.21.1
Madison et al. (2018)[16]	STATsports	Apex / multi-GNSS	18Hz	16 to 20	0.54 ± 0.20 ⇑⇑ (ideal quality)	Apex 2.1.0.4
Massard et al. (2018)[42]	Catapult	MinimaxX S4 / GPS	10Hz	≥ 9	< 1.25 ⇑⇑ (ideal quality)	Sprint 5.1.7
Meylan et al. (2017)[33]	Catapult	Minimax S4 / GPS	10Hz	11.9 ± 1.2	0.96 ± 0.10 ⇑⇑ (ideal quality)	Sprint 5.1.0.1
Park et al. (2019)[38]	Catapult	MinimaxX S4 / GPS	10Hz	8 to 14	>2.0 (no precision)	Sprint 5.1.7
Rago et al. (2019)[32]	QStarz	BT-Q1000 EX / GPS	10Hz	14 ± 1	0.7 ± 0.1 ⇑⇑ (ideal quality)	NR
Ramos et al. (2019)[28]	Catapult	MinimaxX S5/ GPS	10Hz	12.4 ± 0.5	0.75 ± 0.3 ⇑⇑ (ideal quality)	Openfield
Ramos et al. (2017)[27]	Catapult	MinimaxX S5/ GPS	10Hz	15.5 ± 0.5	0.75 ± 0.3 ⇑⇑ (ideal quality)	NR
Scott et al. (2018)[2]	Catapult	Optimeye S5/ GPS	10Hz	12.7 ± 0.7	0.90 ± 0.14 ⇑⇑ (ideal quality)	Openfield 1.14.0
Trewin et al. (2018)[15]	Catapult	MinimaxX S4 / GPS	10Hz	12.1 ± 0.4	0.94 ± 0.04 ⇑⇑ (ideal quality)	Sprint 5.1

NR =Not reported.

Statistical approach

Most studies in their statistical development (Table 8) did not use the logarithmic transformation in its database, which does not bring any problem in using this standard approach to deal with distorted data. The logarithmic transformation can reduce the variability of the data and make data get closer to the normal distribution. If this approach is used, there are limitations when interpreting the relevance of analyzing the transformed data to the original data of interest[21].

Table 8. Statistical approach

Authors	Methods	Software
Abbott et al. (2018)[22]	KS, SW, S, K e 2wA	IBM SPSS Statistics 22
Abbott et al. (2018)[23]	KS, SW, S, K, 2wA e C	IBM SPSS Statistics 22
Abbott et al. (2018)[24]	KS, SW, W, B e C	IBM SPSS Statistics 22
Abbott et al. (2018)[25]	KS, SW, S, K, 2wA, B e C	IBM SPSS Statistics 22
Beato et al. (2019)[29]	SW e C	IBM SPSS Statistics 20
Beato et al. (2018)[40]	BS e t-T	JASP 0.9.1
Curtis et al. (2018)[30]	WT e TK	R Studio 3.2.5
Fitzpatrick et al. (2018)[29]	LR, P e C	NR
Jaspers et al. (2018)[31]	GEE	IBM SPSS Statistics 24
Jones et al. (2019)[34]	LMM e C	IBM SPSS Statistics 22
Knight et al. (2015)[17]	B e C	IBM SPSS Statistics 19
Kyprianou et al. (2019)[41]	TOST	R (3.4.1)
Madison et al. (2018)[16]	2wA, B, t-T e C	IBM SPSS Statistics 22
Massard et al. (2018)[42]	LMM, P, RE e SD	IBM SPSS Statistics 23
Meylan et al. (2017)[33]	CV e P	NR
Park et al. (2019)[38]	KM, GMM, SC, RE e LS	IBM SPSS Statistics 23
Rago et al. (2019)[32]	CV e WS	IBM SPSS Statistics 23
Ramos et al. (2019)[28]	MB e C	NR
Ramos et al. (2017)[27]	MB e C	NR
Scott et al. (2018)[2]	LMM	IBM SPSS Statistics 23
Trewin et al. (2018)[15]	CV e SWC	MS Excel

*Log-Transformed, NR=Not reported, KS=Kolmogorov-Smirnov, SW=Shapiro-Wilk, S=Skewness, K=Kurtosis, 2wA=Two-way ANOVA, C=Cohen's, W=Wilcoxon signed-rank, B=Bonferroni adjustment, BS=Bayesian statistical, t-T=T-test, P=Pearson's correlation, LR=Linear regression, TOST=Two one-sided tests, TK=Turkey pos-hoc, GEE= Generalized estimating equations, LMM=Linear mixed model, WT=Welch test, RE=Random effect, SD=Squared differences post hoc, CV= Coefficient of variation, KM=K-means, GMM=Gaussian mixture model, SC=Spectral Clustering, RE=Random effects, LS=least-squared difference post hoc, WS=Within-subject correlations, MB=Magnitude-based inference and SWC=Smallest worthwhile change

Discussion and implications for practice and research

The present work summarized and analyzed 21 studies related to the load control of football athletes, in which ~ 42% of the selected articles worked with young people <20 years old. Abbott, et al. (2018)[22–25], in their presented 4 articles, the researchers worked with male football athletes all young from the English Premier League, as well as Fitzpatrick et al. (2018)[26]. Ramos et al. (2017)[27] and (2019)[28], used the base of the Brazilian women's team. The risk of bias in these studies was monitored according to the parameters shown in Table 2

and presented in Table 4 with an average of 6.9 points, where only the study by Trewin et al. (2018)[15], did not present average age of the studied group and did not present clear explanations about the context of the presentation of the work, as well as the studies by Madison el al. (2019)[16] and Scott et al. (2018)[2].

Among the 21 studies selected and shown in Table 5, data from the GNSS systems were collected in 1 day of training up to two complete seasons, where ~ 76% of the articles carried out their investigations based on training, with some studies during the competitive period of athletes and ~ 23% during only using official game data, according to Table 5. None of these studies included data from goalkeepers. In addition, the studies by Abbott et al. (2018)[25], Beato et al. (2019)[29], Curtis et al. (2018)[30], Fitzpatrick et al. (2018)[26], Jaspers et al. (2018)[31], Knight et al. (2015)[17], Madison et al. (2019)[16], Rago et al. (2019)[32], Ramos et al. (2017)[27] and Scott et al. (2018)[2]. The studies by Fitzpatrick et al. (2018)[26], Madison et al. (2019)[16] and Scott et al. (2018)[2] used heart rate (HR) records as the only way to control the internal load.

Jaspers et al. (2018)[31] used the perceived effort classification (RPE) using the modified Borg CR-10 scale, and Rago et al. (2019)[32] used the RPE using the Borg CR-100 scale. The studies by Beato (2019)[29], Curtis et al. (2018)[30] and Knight et al. (2015)[17] used both HR and RPE for monitoring, with 1 of these studies also using blood lactate concentration as a form of control. Thus, Curtis et al. (2018)[30] also associated the maximum oxygen volume (VO2 max) as a way to understand the internal changes caused by the exposure of external loads and Abbott et al. (2018)[24] estimated VO2 through field tests to assess maximum aerobic speed (MAS), which is strongly correlated with VO2max.

This study has found three different measures to define the accelerations and decelerations of 2 measures. According to Meylan et al. (2017)[33] and Trewin et al. (2018)[15] in women's football accelerations are those greater than $2.26\ m.s^{-2}$. For Fitzpatrick et al. (2018)[26], the accelerations and decelerations are greater than $2\ m.s^{-2}$ and for Jaspers et al. (2018)[31], Madison et al. (2019)[16] and Ramos et al. (2017, 2019) [27,28] are accelerations greater than $1\ m.s^{-2}$ and decelerations greater than $-1\ m.s^{-2}$. Regarding the variables based on speed, ~ 66% of the studies classify the predominant Sprint variable to understand the intensities of the actions that the athletes are involved. However, the selected studies differ on how to classify and measure this variable. Values of $5.55\ m.s^{-1}$ to $7\ m.s^{-1}\ or\ 25.2\ km.h^{-1}$can be found in the studies.

Two studies present divergent points in their classifications when treating the lowest speeds performed by athletes. Knight et al. (2015)[17] define: stand distance ($< 0.19\ km.h^{-1}$), walk distance ($0.70 - 7.00\ km.h^{-1}\ or\ 0.19 - 1.94\ m.s^{-1}$) jog distance ($7.00 - 14.40\ km.h^{-1}\ or\ 1.94 - 4.00\ m.s^{-1}$), run distance ($14.40 - 20.00\ km.h^{-1} or\ 4.00 - 5.55\ m.s^{-1}$), fast run distance ($20.00 - 23.00\ km.h^{-1} or\ 5.55 - 6.38\ m.s^{-1}$), and Curtis et al. (2018)[30], being: walking distance ($0.00 - 7.19\ km.h^{-1}\ or\ 0.0 - 1.99\ m.s^{-1}$), jogging distance ($7.20 - 14.39\ km.h^{-1}\ or\ 1.99 - 3.99\ m.s^{-1}$), running distance ($14.40 - 21.59\ km.h^{-1} or\ 3.99 - 5.99\ m.s^{-1}$).

Which brings confusion and strangeness when compared to other authors who present other classifications. For Fitzpatrick et al. (2018)[26], speed zones can be divided into distance into distance at very high speed running ($VHSD > 21.00\ km.h^{-1} ou\ 5.83\ m.s^{-1}$), and distance at high speed running ($HSD > 17.00\ km.h^{-1} ou\ 4.72\ m.s^{-1}$). Abbott et al. (2018)[24] classify as VHSD ($5.50 - 7.00\ m.s^{-1}$), HSD ($4.20 - 5.50\ m.s^{-1}$). Jones et al. (2019)[34] bring the classification $<14.40\ km.h^{-1}\ or\ 4.00\ m.s^{-1}$ denominating distance in low-intensity (LSR), as well as Knight et al. (2015)[17] already classified it as below ($14.40 - 19.80\ km.h^{-1}$). Rago et al. (2019)[32] add the moderate zone (distance at moderate-speed running - MSR) with speeds of ($14.40 - 19.80\ km.h^{-1}$), also demonstrated by Jones, et al. (2019)[35] MSR being less than $4.00\ m.s^{-1}$.

The lack of standardization of the measurement units is being presented in some studies in kilometers per hour ($km.h^{-1}$), which does not respect the International Measurement System (SI - *Système international d'unités*), which recommends using meters per second ($m.s^{-1}$) to work with measures related to speed. The SI is a systematized and standardized set of definitions for units of measurement, which aims to standardize and facilitate the measurements and international relations resulting from them. Therefore, the lack of standardization of the data makes it very difficult for most members of the coaching staff to interpret it within the football context[36].

The accuracy of GPS devices has been extensively investigated. Improved accuracy is shown with a sampling frequency of up to 10Hz[14,19] there are no additional benefits in increasing the acquisition frequency[20].

However, the accuracy of the data depends on many factors (atmospheric refraction, multipath, ephemeris, and more), among them, the number of connected satellites and the horizontal precision dilution (HDOP) have been highly discussed in the studies by Malone et al. (2017)[9] and Varley el al. (2017)[14]. However, it is worth mentioning that with the introduction of systems that use more than one constellation of satellites (GPS, GLONASS, Beidou, Compass, Galileo), HDOP or any of the DOPs is not as important as before, since the use of several satellites at the same time generally takes a DOP of less than four[37,38].

Approximately ~ 80% of the studies had an ideal HDOP classification, as shown in Table 7. Only the article by Park et al. (2019)[39]) presents the HDOP description greater than 2.0 and does not relate it to the magnitude of the effect, which makes it difficult to measure whether the data need to be discarded or corrected mathematically, according to the definitions of Moen et al. (1997)[7] and Wu et al. (2012)[40]. The studies by Beato et al. (2018)[41] and (2019)[29] use the HDOP values and number of satellites as the company's standard. Jaspers et al. (2018)[31] reinforce that the values found are in accordance with the study by Malone et al. (2017)[9].

The number of satellites found by the devices is a point to be analyzed, given the geographic location of each study, a variation of greater than five satellites can be verified up to 20 satellites described by the selected studies. Thus, the acquisition of data in most studies was made through the software available with the devices, where only the studies by Ramos et al. (2017 and 2019)[27,28] and Rago et al. (2019)[32] did not show any way of acquiring GPS data. Despite going against the recommendations of Malone et al. (2017)[9] and Varley el al. (2017)[14] it was decided that keeping this study was reasonable because it understands that the data acquisition was made through the product software (Table 7).

The most used software for the applications of the statistical models of the 20 studies was the statistical package for the social sciences (Statistical Package for the Social Sciences), better known as SPSS of the International Business Machines Corporation (IBM). The most common tests were the Kolmogorov-Smirnov and Shapiro-Wilk normality tests, the measures of dispersion of asymmetry and kurtosis, the analysis of bidirectional ANOVA variance and the size of the effect to measure the magnitude of the phenomenon, can be seen in Table 8. The studies by Kyprianou et al. (2019)[42] and Massard et al. (2018)[43] sought to escape from the traditional description to seek equivalence testing, such as the Two one-sided test (TOST) and the grouping techniques widely used in machine learning, respectively.

Study Limitations

It is essential to consider some limitations to this review study. In this paper, 21 studies measure the external load from the use of a GNSS device both in training and in matches. There was no minimum data acquisition period, the description of the number of satellites and the dilution of horizontal precision (HDOP) were significant factors for the

inclusion of studies in football[9,14]. Besides, it does not consider how to acquire acceleration or deceleration data. Either these were obtained through the accelerometer or from the derivation of speed. Thus, it is also limited to the type of activity developed, a period of the season, position in the championship, characteristics of the squad, weather conditions, among other situations that can directly interfere with both male and female athletes.

Conclusion

We can conclude that there is much variation in the applications and methods for collecting and interpreting external load data in football using GNSS technology. Punctually, the lack of classification for uniformity of speed, acceleration, and deceleration thresholds limits the comparisons between the studies, as well as the non-standardization of the measurement units. Several studies have shown that researches classify the zones of speed, acceleration, and deceleration differently and do not use the SI of measures as a way to facilitate the interpretation of the studies. In this way, it is clear that the companies that develop the GNSS devices establish the standard that they find interesting to serve the market, establishing the measurement units they want and not allowing access to raw data, which makes standardization and an intelligent acquisition difficult to process the data.

Also, there is a lack of consistent information about women's football, where only the study by Park et al. (2019)[38], seeks to understand the areas of speed, acceleration, and deceleration for this context. The other studies with female football players use areas established in studies with male football athletes, which often do not meet the studied reality.

It is questionable the lack of inaccuracy in the number of satellites, HDOP, and the software used to acquire data from the GNSS device, where ~ 14% of the selected studies are not transparent or do not present this information. The remaining ~ 85% of the studies presented the average of satellites, the average of HDOP; however, they do not show where this information is obtained.

This review allowed the identification of some gaps in the context of load control in football athletes. The application of load control methods is mainly observed in (1) male and female goalkeepers, (2) in the development of speed, acceleration, and deceleration zones relevant to each studied group (thus not using the standard devices available on the market) and (3) the use of new technologies to correct and improve data from GNSS devices.

Declaration of conflicting interests

The authors declared no potential conflicts of interest with to the research, authorship, and/or publication of this article.

Funding

This work was supported by the Carlos Chagas Filho Foundation for Research Support of the State of Rio de Janeiro under Grant [number E-26/202.638/2018].

References

1. Weston M. Training load monitoring in elite English soccer: a comparison of practices and perceptions between coaches and practitioners. Sci Med Footb 2018; 2: 216–224.
2. Scott D, Lovell R. Individualisation of speed thresholds does not enhance the dose-response determination in football training. J Sports Sci 2018; 36: 1523–1532.

3. Akenhead R, Nassis GP. Training load and player monitoring in high-level football: Current practice and perceptions. Int J Sports Physiol Perform 2016; 11: 587–593.
4. Akenhead R, Harley JA, Tweddle SP. Examining the external training load of an english premier league football team with special reference to acceleration. J Strength Cond Res 2016; 30: 2424–2432.
5. Rico-González M, Los Arcos A, Nakamura FY, et al. The use of technology and sampling frequency to measure variables of tactical positioning in team sports: a systematic review. Res Sport Med 2019; 00: 1–14.
6. Langley RB. Dilution of Precision. GPS World 1999; 10: 52–59.
7. Moen R, Pastor J, Cohen Y. Accuracy of GPS Telemetry Collar Locations with Differential Correction. J Wildl Manage 1997; 61: 530.
8. Vaz JA, Pissardini RDS, Fonseca Junior ES Da. Comparação da cobertura e acurácia entre os sistemas GLONASS e GPS obtidas dos dados de observação de uma estação da Rede Brasileira de Monitoramento Contínuo. Rev Bras Cartogr 2013; 65: 529–539.
9. Malone JJ, Lovell R, Varley MC, et al. Unpacking the black box: Applications and considerations for using gps devices in sport. Int J Sports Physiol Perform 2017; 12: 18–26.
10. Rago V, Brito J, Figueiredo P, et al. Methods to collect and interpret external training load using microtechnology incorporating GPS in professional football: a systematic review. Res Sport Med 2019; 00: 1–22.
11. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: Explanation and elaboration. PLoS Med; 6. Epub ahead of print 2009. DOI: 10.1371/journal.pmed.1000100.
12. Saw AE, Main LC, Gastin PB. Monitoring the athlete training response : subjective self-reported measures trump commonly used objective measures : a systematic review. Br J Sports Med 2016; 50: 281–291.
13. Horsley T, Dingwall O, Tetzlaff JM, et al. Checking reference lists to find additional studies for systematic reviews. Cochrane Database Syst Rev. Epub ahead of print 2009. DOI: 10.1002/14651858.MR000026.
14. Varley MC, Jaspers A, Helsen WF, et al. Methodological considerations when quantifying high-intensity efforts in team sport using global positioning system technology. Int J Sports Physiol Perform 2017; 12: 1059–1068.
15. Trewin J, Meylan C, Varley MC, et al. The match-to-match variation of match-running in elite female soccer. J Sci Med Sport 2018; 21: 196–201.
16. Madison G, Patterson SD, Read P, et al. Effects of small-sided game variation on changes in hamstring strength. J Strength Cond Res 2019; 33: 839–845.
17. Knight TJ, Scott BR, Jalilvand F, et al. The effect of changing player numbers on the physiological responses and time-motion characteristics of a soccer-specific training drill. Int J Perform Anal Sport 2015; 15: 452–470.
18. Varley MC, Fairweather IH, Aughey RJ. Validity and reliability of GPS for measuring instantaneous velocity during acceleration, deceleration, and constant motion. J Sports Sci 2012; 30: 121–127.
19. Rampinini E, Alberti G, Fiorenza M, et al. Accuracy of GPS devices for measuring high-intensity running in field-based team sports. Int J Sports Med 2015; 36: 49–53.
20. Scott M, Scott T, Kelly V. The Validity and Reliability of Global Positioning Systems in Team Sport: A Brief Review. J Strength Cond Res 2016; 30: 1470–1490.
21. Feng C, Wang H, Lu N, et al. Log-transformation and its implications for data analysis. Shanghai Arch Psychiatry 2014; 26: 105–109.

22. Abbott W, Brickley G, Smeeton NJ. Positional differences in GPS outputs and perceived exertion during soccer training games and competition. 2018. Epub ahead of print 2018. DOI: 10.1519/JSC.0000000000002387.
23. Abbott W, Brickley G, Smeeton NJ, et al. Individualizing acceleration in english premier league academy soccer players. J Strength Cond Res 2018; 32: 3503–3510.
24. Abbott W, Brickley G, Smeeton NJ. An individual approach to monitoring locomotive training load in English Premier League academy soccer players. Int J Sport Sci Coach 2018; 13: 421–428.
25. Abbott W, Brickley G, Smeeton NJ. Physical demands of playing position within English Premier League academy soccer. J Hum Sport Exerc 2018; 13: 285–295.
26. Fitzpatrick JF, Hicks KM, Hayes PR. Dose-response relationship between training load and changes in aerobic fitness in professional youth soccer players. Int J Sports Physiol Perform 2018; 13: 1365–1370.
27. Ramos GP, Nakamura FY, Pereira LA, et al. Movement Patterns of a U-20 National Women's Soccer Team during Competitive Matches: Influence of Playing Position and Performance in the First Half. Int J Sports Med 2017; 38: 747–754.
28. Ramos GP, Nakamura FY, Penna EM, et al. Activity Profiles in U17, U20, and Senior Women's Brazilian National Soccer Teams During International Competitions: Are There Meaningful Differences? J strength Cond Res 2019; 33: 3414–3422.
29. Beato M, De Keijzer KL, Carty B, et al. Monitoring fatigue during intermittent exercise with accelerometer-derived metrics. Front Physiol; 10. Epub ahead of print 2019. DOI: 10.3389/fphys.2019.00780.
30. Curtis R, Huggins R, Looney D, et al. Match Demands of National Collegiate Athletic Association Division I Men's Soccer. J Strength Cond Res 2018; 32: 2907–2917.
31. Jaspers A, Kuyvenhoven JP, Staes F, et al. Examination of the external and internal load indicators' association with overuse injuries in professional soccer players. J Sci Med Sport 2018; 21: 579–585.
32. Rago V, Krustrup P, Martín-Acero R, et al. Training load and submaximal heart rate testing throughout a competitive period in a top-level male football team. J Sports Sci. Epub ahead of print 2019. DOI: 10.1080/02640414.2019.1618534.
33. Meylan C, Trewin J, McKean K. Quantifying explosive actions in international women's soccer. Int J Sports Physiol Perform 2017; 12: 310–315.
34. Jones RN, Greig M, Mawéné Y, et al. The influence of short-term fixture congestion on position specific match running performance and external loading patterns in English professional soccer. J Sports Sci 2019; 37: 1338–1346.
35. Jones RN, Greig M, Mawéné Y, et al. The influence of short-term fixture congestion on position specific match running performance and external loading patterns in English professional soccer. J Sports Sci 2019; 37: 1338–1346.
36. Gauthier-Villars. Comptes rendus des séances de la onzième conférence générale des poids et mesures. Paris, 1960.
37. Jyothirmaye S, Srinivas VS, Ramu B. Fast Satellite Selection Techniques and DOPs for Multi-GNSS Positioning S. 2019 URSI Asia-Pacific Radio Sci Conf 2019; 1–4.
38. Zhang Q, Chen Z, Rong F, et al. Preliminary Availability Assessment of Multi-GNSS: A Global Scale Analysis. IEEE Access 2019; 7: 146813–146820.
39. Park LAF, Scott D, Lovell R. Velocity zone classification in elite women's football: where do we draw the lines? Sci Med Footb 2019; 3: 21–28.
40. Wu S, Liu Y, Liu L, et al. The Distributions of HDOP and VDOP in GNSS and a Corresponding New Algorithm of Fast Selecting Satellites. In: China Satellite Navigation Conference (CSNC) 2012 Proceedings, pp. 445–453.

41. Beato M, Coratella G, Stiff A, et al. The validity and between-unit variability of GNSS units (STATSports apex 10 and 18 Hz) for measuring distance and peak speed in team sports. Front Physiol; 9. Epub ahead of print 2018. DOI: 10.3389/fphys.2018.01288.
42. Kyprianou E, Lolli L, Haddad H Al, et al. A novel approach to assessing validity in sports performance research: integrating expert practitioner opinion into the statistical analysis. Sci Med Footb; 3. Epub ahead of print 2019. DOI: 10.1080/24733938.2019.1617433.
43. Massard T, Eggers T, Lovell R. Peak speed determination in football: is sprint testing necessary? Sci Med Footb 2018; 2: 123–126.

3 ESTUDO 3- GPS TECHNOLOGY TO CONTROL OF EXTERNAL DEMANDS OF ELITE BRAZILIAN FEMALE FOOTBALL PLAYERS DURING COMPETITIONS

Abstract

In order to understand the performance of female football players in concomitantly competitions. **Purpose:** This study aimed to determine the behavior of the external load demands of elite Brazilian soccer players by position and playing time, using data from global positioning technology (GPS) at a data acquisition frequency of 10Hz from a Polar Team Pro. **Methods:** Twenty-three professional players from a top-flight Brazilian women's soccer team participated in the study. The women were 27.65 ± 4.66 years, 165.35 ± 5.82 cm and 60.91 ± 5.34 kg. Data were collected during two competitions disputed simultaneously, the Campeonato Paulista (The São Paulo State Championship) and the Campeonato Brasileiro (Brazilian Championship) A1 series of 2019. Nine variables related to the external locomotor demand performed by the players in the disputed matches were analyzed. The statistical procedures used were the two-way ANOVA with Bonferroni post hoc for data adjustment, and Cohen's effect size (d), with a p- value<0.05 adopted to demonstrate statistical significance. **Results:** The significant differences in the variables (total distance, WD, JD, RD, SD, A1, A2, D1, and D2) were analyzed, and the results showed a decrease in the external locomotor demand of the loads during the matches played by these athletes. **Conclusion:** These results should be used to improve load control and adapt training practices in women's soccer teams.

Keywords: Load Control, External Load Control, Female Football, Global Navigation Satellite System

Introduction

Understanding the nature of the effect a football match has on athletes is essential to optimizing performance and thus being able to develop and implement effective interventions through the quantification of athletes' workloads (Weston, 2018). Global positioning system (GPS) technology can improve understanding of this phenomenon, when used to track activity on the playing field in open-air team sports (Cummins et al., 2013).

In 2015, the International Council of the Football Association sanctioned the use of wearable technologies during official matches, or competitions. This made it possible to use small portable GPS units simultaneously with heart rate (HR) monitoring to characterize the external load and internal work performed (Bourdon et al., 2017), as an integral part of the daily life of football clubs (Akenhead & Nassis, 2016).

Initially, the implemented GPS technology was used exclusively for military or scientific purposes. However, with its advancement, more extensive applications have emerged, such as use in team sports (Beato et al., 2018; Scott et al., 2016). Thus, essential components can be quantified based on information from this technology, such as distances covered, speeds, accelerations, and decelerations (Akenhead et al., 2016), or even movement pattern and physical impact of training sessions or games (Hennessy & Jeffreys, 2018).

Despite its practical applications, the significant technological advance, and the massive amount of research using portable GPS devices (Rago et al., 2019), certain problems have been found when they are used, these : (1) the reliability and validity of the device; (2) the form of data collection and processing; (3) the satellite connection and its accuracy; and (4) data exclusion criteria (Malone et al., (2017) and Varley et al., (2017)).

Thus, it is feasible that professionals are more concerned with exploring the data collected from these devices, always aiming at a more sophisticated analysis, rather than understanding the validity and reliability of the system used. Often, these professionals are conditioned by the market before essential information about the accuracy of devices is known (Russell et al., 2016).

The use of GPS in professional football can facilitate the collection of variables necessary to control both the internal and external load of players during training and / or games (Buchheit et al., 2015). The quantification of training and game load can help decrease injuries and improve athletes' performance in team sports (Cummins et al., 2013). This may indicate that multiple objective external and internal load variables should be included to accurately describe the physical and physiological demands of team sports in general and football in particular.

Football has experienced a significant increase in the number of players participating in international competitions, professional leagues, and leisure leagues (Marqués-Jiménez et al., 2017). The physical and physiological loads of male football players, within the competitive context at different ages and levels, have been extensively examined in the literature for an extended period (Carling et al., 2008). On the other hand, even with the increase in investigations within women's football in recent years, there is little data available on such demands on female players (Datson et al., 2014).

GPS studies to quantify these loads in international women's football show higher HR values and more prolonged duration at high intensity during the game compared with training, which indicates higher demands on the players' internal load during the game (Ohlsson et al., 2015).

A successful assessment of the physical loads imposed on football players requires an accurate assessment of both internal and external demands (Gaudino et al., 2015). Studies using GPS devices have shown that professional football players cover distances of

approximately 8 ± 3 km during matches with a range of up to 119 meters per minute ($m.min^{-1}$) (Bradley et al., 2014; Trewin et al., 2018).

A study by Andersson et al., (2010), shows that high-intensity running represents up to 12% of the total distance. Midfield players tend to cover a longer distance, while defenders cover a shorter distance, both in total area covered and in high intensity running.

However, the distance covered at different speed ranges should not be the only consideration for assessing an athlete's workload. Specific low-intensity movements in football such as a) accelerations, b) decelerations, c) side and back runs, and even d) passes, are often neglected, leading to an erroneous estimation of the workloads performed by the players (Reche- Soto et al., 2019).

The workload of male and female players also differs between game positions, with midfield players achieving a higher load when compared to defenders or forwards (Park et al., 2019; Reche- Soto et al., 2019; Strauss et al., 2019).

When analyzing the physiological demands of female football athletes, when divided by activity sectors on the field of play, it is noticed that athletes who act on the sides of the pitch have an increase in distance covered during the game, but do not necessarily have a difference in aerobic conditioning parameters, despite their futsal counterparts, the wings, having higher resistance to fatigue (Alvares et al., 2017; Perrier-Melo et al., 2018). This aspect differs from men's football, where a significant difference is perceived regarding maximum oxygen consumption when the squad is divided by game positions (Barbalho et al., 2017).

In addition to the physiological capacities, another essential point for players' athletic performance is decision-making ability, since this variable is an intervening factor for the quantification of data related to external load demands. In athletes of both genders, positioning and age influence decision-making style on the field. Those who are multipurpose, that is, who work in more than one position, mainly in the game transition zone, are able to obtain better decision rates (Teixeira, 2019).

Player load also differs between positions in a football match, with midfielders reaching a higher load compared with forwards and defenders, thereby generally achieving a significantly higher relative total distance than that observed for defenders, and consequently, with a higher chance of fatigue (Dalen et al., 2016).

Forwards act in the area of the field where more significant activity at high intensity is necessary, such as successive sprints and greater decision-making power. Improving the players' ability to perform these activities should be emphasized in aerobic and resistance training programs, developing in the players a higher possibility of overcoming opponents, improving decision-making power, and creating more goal opportunities (Strauss et al., 2019).

The physiological stress experienced by players can result in fatigue that leads to loss of performance (Marqués-Jiménez et al., 2017), as such, performance is associated with appropriate recovery between matches and training (Andersson et al., 2008).

Therefore, detailing the workload performed by players during a competitive match provides a more exceptional view of the physical demands in women's football, highlighting the characteristics related to performance and providing information to coaches and members of the coaching staff for more efficient and precise load control of each athlete (Strauss et al., 2019).

Thus, the aim of this study was to determine the behavior of external load demands of elite female players in Brazilian football, using data on position and playing time in two official competitions running simultaneously.

Material and Methods

Subjects

Twenty-three players from a top-flight professional team in Brazilian women's football participated in this study (mean ± standard deviation, age: 27.65 ± 4.66 years, height: 165.35 ± 5.82 cm, weight: 60.91 ± 5.34 kg). All players compete at national level, displaying experience both in-game and in training. The players were also classified according to their playing position: Defenders (DF): 30.29 ± 4.89 years; 163.75 ± 5.74 cm; 59.79 ± 6.53 kg; Midfielders (MF): 27.56 ± 4.48 years; 163.76 ± 3.17 cm; 60.28 ± 3.08 kg; and Forwards (FW): 25.14 ± 3.67 years; 166.28 ± 5.19 cm; 61.25 ± 5.87 kg.

The Human Research Ethics Committee of the Pedro Ernesto University Hospital (HUPE) of the State University of Rio de Janeiro (UERJ) approved the study (10529119.8.0000.5259) and the participants were informed of the possible benefits and risks of the investigation before signing an institutionally approved informed consent form to participate in the study.

Procedures

All matches were held on pitches with official FIFA® dimensions (100 × 75 m), during 12 matches of the 2019 São Paulo State Championship and the last 11 matches of the 2019 A1 Brazilian Championship. The profiles on the players' locomotor activities in the matches were obtained using global positioning system (GPS) units with a frequency of 10Hz (Polar Team Pro™, Polar Electro, Kempele, Finland), as were the physiological measurements on the number of heart beats per minute and their percentage, which were also measured using the same device.

External load measures

The external loads during the matches of the first phase of the São Paulo State Championship and the Brazilian women's championship were acquired using a 10Hz GPS sensor, recorded by a microelectromechanical sensor (MEMS) (Polar Team Pro™, Polar Electro, Kempele, Finland). Each player was equipped with a GPS unit, coupled to a 200Hz MEMS motion sensor (tri-axial accelerometer, gyroscope, and magnetometer) (Polar Team Pro™, Polar Electro, Kempele, Finland). The monitoring device (39 g, 36 × 68 × 13 mm) was located on the chest, attached by a specific strap around the chest, with the sensor positioned over the entire xiphoid process.

GPS devices have shown increasingly improved accuracy at a sampling frequency of up to 10Hz (Malone et al., 2017), with no additional data-acquisition benefits from increasing the frequency (Scott et al., 2016). All devices were turned on 15 to 30 minutes before data collection to enable the acquisition of satellite signals and synchronize the device's clock with the satellite's atomic clock (Maddison & Ni Mhurchu, 2009). This procedure enabled the connection and acquisition of more than four satellites and a horizontal dilution of precision (HDOP) of less than 5, according to information from the manufacturer. The same unit was also used by each player in all matches to reduce the measurement error between units. The reliability of similar devices has been well documented in the literature (Clemente et al., 2019; Evangelio et al., 2019; Silva et al., 2015).

Polar Team Pro™ hardware and software were used to objectively record the mechanical aspects related to each player. Subsequently, the data were processed through the

Polar Team Pro™ device web platform and exported in the text format regulated by RFC 4180 (comma-separated values - CSV).

Game Analysis

Activity demands were measured using portable GPS devices (10 Hz, Polar Team Pro™, Polar Electro, Kempele, Finland) and the speed zones defined by the manufacturer for the football modality were used to categorize the distance the player remained active (Zone 1: $1.66\ to\ 3.33\ m.s^{-1}$; Zone 2: $3.34\ to\ 4.44\ m.s^{-1}$; Zone 3: $4.45\ to\ 5.55\ m.s^{-1}$; Zone 4: $5.56\ to\ 6.38\ m.s^{-1}$; Zone 5: $>\ 6.39\ m.s^{-1}$). These speed-related zones were transformed from $km.h^{-1}$ to $m.s^{-1}$ respecting that recommended by the International Measurement System (Système International d'Unités) (Gauthier-Villars, 1960), with the purpose of facilitating the interpretation of the data presented.

The acceleration and deceleration frequencies were measured using the accelerometer present in the Polar Team system with 200 Hz sampling (Polar Team Pro™, Polar Electro, Kempele, Finland). The zones defined as standard by the manufacturer were also used to categorize acceleration intensities (low: $0.50\ to\ 0.99\ m.s^{-2}$; moderate: $1.00\ to\ 1.99\ m.s^{-2}$; high: $2.00\ to\ 2.99\ m.s^{-2}$; very high: $>3.00\ m.s^{-2}$) and deceleration intensities (low: $-0.99\ to\ -0.50\ m.s^{-2}$; moderate: $-1.99\ to\ 1.00\ m.s^{-2}$; high: $-2.99\ to\ -2.00\ m.s^{-2}$; very high: $<-3.00\ m.s^{-2}$).

Therefore, for analysis purposes, the activities performed in the matches were divided into the following categories: walking distance (WD): $\leq 3.33\ m.s^{-1}$), jogging distance (JD): $3.34\ to\ 4.44\ m.s^{-1}$, running distance (RD): $4.45\ to\ 5.55\ m.s^{-1}$, sprinting distance (SD): $\geq 5{,}55\ m.s^{-1}$, acceleration number (Acc) $\geq 2.00\ m.s^{-1}$ and (IV) deceleration number (Des) $\leq -2.00\ m.s^{-1}$ (A Strauss et al., 2019).

Statistical Analysis

The data were analyzed using the IBM SPSS Statistics 23 statistical program and presented as mean and standard deviation. The normality and homogeneity of the variance of the sample data were analyzed using the Kolmogorov-Smirnov and Levene tests. Bartlett's test was used to verify the sphericity of the data. Two-way ANOVA was used, having the position of the independent variables and playing time as factors, followed by Bonferroni's post hoc with data adjustment to identify possible differences in the study variables. Cohen's effect size (d) (Cohen, 1988), was calculated to analyze the effect's magnitude. The magnitudes were interpreted qualitatively using the limits: insignificant (<0.2); small ($0.2-0.6$); moderate ($0.6-1.2$); large ($1.2-2.0$); very large ($2.0-4.0$); and almost perfect (>4.0) (Hopkins et al., 2009). The value of $p<0.05$ was adopted for statistical significance.

Results

The physical variables of the matches for the external demands described between the different positions are shown in Table 1. It can be seen that for each position on the field of play, different variables have no normality and homogeneity according to the Kolmogorov-Smirnov test. Thus, it is observed that total distance, walking distance, acceleration in zone 2, that is, from $1.00\ to\ 1.99\ m.s^{-1}$, and deceleration in zone 2 ($-1.00\ to\ -1.99\ m.s^{-1}$), are the only variables that demonstrate absence of normality and homogeneity in all three positions of the playing field. To demonstrate this behavior, analysis of variance (ANOVA) with Bonferroni's adjustment was used to treat the collected data.

Table 1. Description of external load variables

Variable	Defenders	Midfielders	Forwards
Total distance (m)	4,085.63 ± 1,091.73*	4,080.62 ± 1,351.94*	3,648.74 ± 1,323.96*
WD (m)	1,496,23 ± 424.62*	1,563.36 ± 568.72*	1,379.82 ± 542.88*
JD (m)	711.25 ± 240.89	745.60 ± 339.98	581.63 ± 261.21
RD (m)	302.49 ± 109.38	323.61 ± 148.55	273.47 ± 126.95
SD (m)	149.96 ± 80.27	136.63 ± 78,15*	166.56 ± 110.72*
A1 ($m.s^{-1}$)	125.33 ± 37.93*	123.63 ± 46.54*	105.95 ± 40.68*
A2 ($m.s^{-1}$)	38.22 ± 12.63*	36.21 ± 15.53	35.01 ± 14.82
D1 ($m.s^{-1}$)	40.61 ± 14.13*	39.75 ± 16.88	35.94 ± 15.57*
D2 ($m.s^{-1}$)	127.20 ± 39.55*	123.08 ± 46.75*	106.73 ± 42.00*

*lack of normality according to Kolmogorov-Smirnov test values, $p < 0.05$, WD = walking distance ($\leq 3.33\ m.s^{-1}$), JD = jogging distance: ($3.34 - 4.44\ m.s^{-1}$), RD = running distance ($4.45 - 5.55\ m.s^{-1}$), SD = Sprinting distance ($\geq 5.55\ m.s^{-1}$), A2 = acceleration zone 2 ($1.00\ a\ 1.99\ m.s^{-1}$), A1 = acceleration zone 1 ($\geq 2.00\ m.s^{-1}$), D2 = deceleration zone 2 ($-1.00\ a - 1.99\ m.s^{-1}$), D1 = deceleration zone 1 ($\leq -2.00\ m.s^{-1}$).

Table 2 shows the external load demands during official games divided into two groups, these being the 1st and 2nd half of matches for the different game positions described in this study. Thus, we highlight the statistical difference between each position and the magnitude of the effect size (d) presented, as well as the 95% confidence interval for d.

When comparing Defenders and Forwards, the magnitude of the effect size was moderate for jogging distance, acceleration zone 2, and deceleration zone 2. In the comparison between Midfielders and Forwards, the moderate effect size appears only for acceleration zone 1, even though it does not show any significant difference by the method used. It is also noteworthy that when comparing Defenders and Midfielders, the effect size is approximately 89% insignificant in all the studied samples.

Table 2. Load activity of different positions in 45 minute periods

Variable	Time	Defenders (DF)	Midfielders (MF)	Forwards (FW)	Cohen (*d*) DF → MF	Cohen (*d*) DF → FW	Cohen (*d*) MF → FW
Total Distance (m)	1º	4,347.23 ± 1,123.64	4,596.21 ± 1195.95	4,464,50 ± 1040.26	-0.214 (Insignificant effect) CI95% [-0.51; 0.09]	-0.108 (Insignificant effect) CI95% [-0.44; 0.23]	-0.669 (Insignificant effect) CI95% [-0.99; 0.34]
	2º	3,859.29 ± 1,016.21 ∂	3,647.15 ± 1327.62 ρ	3,137,86 ± 1226.21	0.177 (Insignificant effect) CI95% [-0.10; 0.45]	0.637 (Insignificant effect) CI95% [0.34; 0.93]	1.205 (Moderate effect) CI95% [0.91; 1.49]
WD (m)	1º	1,609,44 ± 449.71	1,785,37 ± 520.64	1,707,32 ± 475.54	-0.359 (Insignificant effect) CI95% [-0.66; -0.05]	-0.212 (Insignificant effect) CI95% [-0.55; 0.12]	0.155 (Insignificant effect) CI95% [-0.17; 0.47]
	2º	1,398,28 ± 377.47 σ	1,376.71 ± 541.61 ρ	1,174,72 ± 479.64	0.045 (Insignificant effect) CI95% [-0.23; 0.32]	0.515 (Small effect) CI95% [0.22; 0.80]	0.393 (Small effect) CI95% [0.12; 0.66]
JD (m)	1º	764.03 ± 259.51	849.25 ± 336.94 ρ	716.97 ± 266.12	-0.280 (Insignificant effect) CI95% [-0.58; 0.02]	0.179 (Insignificant effect) CI95% [-0.16; 0.51]	0.000 (Insignificant effect) CI95% [-0.32; 0.32]
	2º	665.58 ± 214.71	658.45 ± 318.70 ρ	496.87± 220.15	0.026 (Insignificant effect) CI95% [-0.25; 0.30]	0.775 (Moderate effect) CI95% [0.31; 0.86]	0.583 (Small effect) CI95% [0.31; 0.86]
RD (m)	1º	319.52 ± 115.65 ∂	368.02 ± 150.37	326.56 ± 130.38	-0.357 (Insignificant effect) CI95% [-0.66; -0.05]	-0.058 (Insignificant effect) CI95% [-0.39; 0.28]	0.290 (Small effect) CI95% [-0.03; 0.61]
	2º	287.76 ± 102.03	286.27 ± 136.85 ρ	240.21 ± 113.26	0.012 (Insignificant effect) CI95% [-0.27; 0.29]	0.440 (Small effect) CI95% [0.15; 0.73]	0.364 (Small effect) CI95% [0.09; 0.64]
SD (m)	1º	163.96 ± 83.29	157.74 ± 88.26 ρ	197.32 ± 113.67	0.072 (Insignificant effect) CI95% [-0.23; 0.37]	-0.340 (Insignificant effect) CI95% [-0.68; 0.00]	-0.400 (Insignificant effect) CI95% [-0.72; -0.07]
	2º	137.84 ± 75.97	118.89 ± 63.68	147.29 ± 104.88	0.273 (Small effect) CI95% [-0.01; 0.55]	-0.102 (Insignificant effect) CI95% [-0.39; 0.18]	-0.332 (Insignificant effect) CI95% [-0.60; -0.06]
A1 ($m.s^{-1}$)	1º	132.84 ± 38.30	138.54 ± 43.09	127.68 ± 36.09	-0.139 (Insignificant effect) CI95% [-0.44; 0.16]	0.138 (Insignificant effect) CI95% [-0.20; 0.47]	0.268 (Small effect) CI95% [-0.05; 0.59]
	2º	118.82 ± 36.58 σ	111.10 ± 45.80 ρ	92.34 ± 37.49	0.184 (Insignificant effect) CI95% [-0.10; 0.46]	0.714 (Moderate effect) CI95% [0.42; 1.01]	0.445 (Small effect) CI95% [0.17; 0.72]
A2 ($m.s^{-1}$)	1º	41.08 ± 12.89	41.08 ± 15.09	41.34 ± 14.31	0.000 (Insignificant effect) CI95% [-0.30; 0.30]	-0.019 (Insignificant effect) CI95% [-0.35; 0.32]	-0.018 (Insignificant effect) CI95% [-0.34; 0.30]
	2º	35.75 ± 11.92	35.75 ± 14.74	31.05 ± 13.79	0.000 (Insignificant effect) CI95% [-0.28; 0.28]	0.362 (Small effect) CI95% [0.07; 0.65]	0.692 (Moderate effect) CI95% [0.41; 0.97]
D1 ($m.s^{-1}$)	1º	44.05 ± 13.51	46.65 ± 16.29	44.16 ± 14.78	-0.008 (Insignificant effect) CI95% [-0.34; 0.33]	-0.172 (Insignificant effect) CI95% [-0.47; 0.13]	0.158 (Insignificant effect) CI95% [-0.16; 0.48]
	2º	37.63 ± 14.04 σ	33.96 ± 15.15	30.79 ± 13.80	0.250 (Small effect) CI95% [-0.03; 0.53]	0.492 (Small effect) CI95% [0.20; 0.78]	0.218 (Small effect) CI95% [-0.05; 0.49]
D2 ($m.s^{-1}$)	1º	134.13 ± 38.96	135.62 ± 42.31	128.98 ± 36.12	-0.036 (Insignificant effect) CI95% [-0.34; 0.26]	0.137 (Insignificant effect) CI95% [-0.20; 0.47]	0.166 (Insignificant effect) CI95% [-0.16; 0.49]
	2º	121.20 ± 39.28 σ	112.54 ± 47.86 ρ	92.80 ± 39.47	0.026 (Insignificant effect) CI95% [-0.25; 0.30]	0.721 (Moderate effect) CI95% [0.42; 1.01]	0.394 (Small effect) CI95% [0.12; 0.66]

∂ Difference between Defenders and Midfielders, σ Difference between Defenders and Forwards, ρ Difference between Midfielders and Forwards, WD = walking distance ($\leq 3.33\, m.s^{-1}$), JD = jogging distance: ($3.34 - 4.44\, m.s^{-1}$), RD = running distance ($4.45 - 5.55\, m.s^{-1}$), SD = sprinting distance ($\geq 5.55\, m.s^{-1}$), A2 = acceleration zone 2 ($1.00\ to\ 1.99\, m.s^{-1}$), A1 = acceleration zone 1 ($\geq 2.00\, m.s^{-1}$), D2 = deceleration zone 2 ($-1.00\ to\ -1.99\, m.s^{-1}$), D1 = deceleration zone 1 ($\leq -2.00\, m.s^{-1}$).

Discussion

The external physical locomotor demands demonstrated through the acquisition of GPS data in 23 official matches showed that the defenders covered a greater total distance with a difference of approximately 5.01 m to midfielders. This characteristic of the analyzed team is different from the findings of Strauss et al. (2019), where midfield players cover a greater distance in matches, as previously shown by Andersson et al. (2010) in international players.

When observing the absolute values of total distance, it can be seen that the players present in this study had values in units of thousands (1,000), as per the studies by Anita Strauss et al. (2019), using GPS with 30 sub-elite athletes (age: 22.8 ± 2.4). This differs from the 17 international level athletes (age: 27.0 ± 1.0) from Sweden and Denmark in the study by Andersson et al. (2010) using video analysis, where the absolute values are described in units of ten thousands (10,000).

The study by Ramos et al. (2017) with U20 athletes from the Brazilian women's national team, presents a scale of values similar to that found, for example, in the average total distance of the defenders, with 8,201.60 m and this study being 8,206.52 m. On the other hand, when the data are compared with a senior Brazilian female team, as described in the study by Ramos et al. (2019), defenders have an average of 10,003.40 m in absolute values, which is similar to that found by Andersson et al. (2010), also in senior teams. Thus, from the use of other variables for a more in-depth analysis, differences between teams and competitions can be identified around the world.

The players who act in defense, in midfield, and in attack cover average distances (4,085.63m; 4,080.62m; 3,648.74m) greater than those presented in the studies by Dalen et al., (2016), Park et al., (2019), Reche- Soto et al., (2019), and Anita Strauss et al., (2019). This shows that total distance is a variable relative to the match and the opponent faced in each competition, requiring variables that determine the context for purposes of comparison.

Total distance is a variable present in several studies that use GPS technology to measure external load in football players of both genders, as documented by Rago et al. (2019). In the present study, in Table 2, the total distance only showed significant results in the second half of the matches, when comparing athletes in the defensive and attacking positions ($p < 0.001$). Therefore, understanding the game model of the teams to promote a better definition of the distribution of the players on the field of play can bring an improvement to the data analysis of athletes who work in the midfield and the attack.

Likewise, generic speed thresholds are universally used in monitoring external loads in different sporting contexts (Akenhead & Nassis, 2016). This evaluation enables determination of load evolution in training and games, providing comparisons between athletes and teams.

The activity demands for the speed zones showed a significant difference in the 1st half of the matches between midfielders and forwards for the variables JD ($p = 0.010$) and SD ($p = 0.018$); and between defenders and midfielders for RD ($p = 0.038$).

In the 2nd half, the variables JD and RD between midfielders and forwards show $p < 0.001$ and $p = 0.025$, respectively, but with a small magnitude of effect. Thus, it can be understood that these isolated variables cannot describe sports activity during match times, as found by Anita Strauss et al. (2019).

In order to understand the influence of acceleration and consequent deceleration actions within a game, and whether a given acceleration action will generate an opposite deceleration action of the same quantity and magnitude, actions were categorized into two zones, acceleration zone 1 (A1) with $\geq 2.00\, m.s^{-1}$ and zone 2 (A2) with

$1.00\ to\ 1.99\ m.s^{-1}$, and deceleration zone 1 (D1) with values from $-1.00\ to -1.99\ m.s^{-1}$ and deceleration zone 2 (D2) with values $\leq -2.00\ m.s^{-1}$. according to G P Ramos et al. (2017) and Joshua Trewin et al. (2017).

The acceleration variable in zone 2 does not show any significant difference in any of the positions in either half of the game, while deceleration in zone 1 shows a difference in the actions between Defenders and Forwards ($p = 0.004$) in the second half of the game. However, the actions related to acceleration zone 1 and deceleration zone 2 reflected the significance of actions between defenders and forwards ($p < 0.001$) and between midfielders and forwards ($p = 0.004$) in the second half of the matches, as per studies by Guilherme P. Ramos et al. (2019) and Joshua Trewin et al. (2017).

In the game model presented, which is characterized by progression up the field of play with the maintenance of ball possession, defenders presented a highly significant number of accelerations and decelerations, with an average of 38.22 accelerations above $\geq 2.00\ m.s^{-1}$ and 40.61 decelerations below $\leq -2{,}00\ m.s^{-1}$, showing a more significant number of these actions than other athletes from other positions. This may be related to the support of the flanks that were included in this category for study purposes and are players who constantly need to accelerate and decelerate during a match.

Conclusion

Generally, there is an overall decrease in the external locomotor demand for loads during matches with professional Brazilian female football players. These differences were observed using playing field positions and analyzed according to the behavior of each variable during two simultaneous competitions.

Based on the results, there is a need to improve the components of physical fitness specific to the games of these players, especially when there is a decrease in the means of the variables among all positions during the different periods of the matches in the competitions. This decrease presents moderate magnitudes between defenders and midfielders, and midfielders and forwards, which may be associated with physiological deterioration of the players during game halves and competitions being played concurrently.

Thus, frequent evaluation of the players where it is possible to measure components such as a) concentric and eccentric strength of the muscles of the lower limbs, and b) tests that can simulate agility, speed, change of direction, and decision making of the athletes, are essential for the deployment of more assertive algorithms from data analysis. This greater assertiveness using data from external loads from GPS and tests organized throughout the season represents a different view to improving the components of these players' physical fitness.

This information is relevant so that coaches and technical staff can discern the physical standards for each position. The study showed that there are significant differences between the positions of defense, midfield, and attack among professional female football players. The conclusions of this study provided a theoretical justification to explain the level of performance achieved by the players, which can be manipulated by tasks in training that provide relevant implications for improving physical interactions during matches.

Also, the use of micro-sensor technology that makes it possible to quantify the external locomotor demands of athletes through metrics originating from global satellite positioning (GPS) is suggested based on the variability of the observed measures. However, it is advisable for members of the coaching staff to examine their players' individual variation whenever possible, in order to make inferences regarding changes in performance between the observed matches. Therefore, we understand the lack of transparency of the company that supplies the device as a limitation of this study, as the company does not present the number

of satellites or the horizontal precision value (HDOP) at the time of data acquisition; also, how the positions of defense, midfield and attack are classified is a further limitation.

Declaration of conflict of interests

The authors declare no potential conflicts of interest with the research, authorship, and/or publication of this article.

Funding

This work was supported by the Carlos Chagas Filho Foundation for Research Support of the State of Rio de Janeiro under Grant [number E-26/202.638/2018].

References

Akenhead, R, & Nassis, G. P. (2016). Training load and player monitoring in high-level football: Current practice and perceptions. *International Journal of Sports Physiology and Performance*, *11*(5), 587–593.

Akenhead, Richard, Harley, J. A., & Tweddle, S. P. (2016). Examining the external training load of an english premier league football team with special reference to acceleration. *Journal of Strength and Conditioning Research*, *30*(9), 2424–2432.

Alvares, P. D., Diniz, R. R., Santana, P. V. A., Lima, F. A., Costa, L. L. da, Leite, R. D., Carvalho, R. P. de, Reis, A. D., & Pires, F. de O. (2017). Potência anaeróbia máxima e índice de fadiga em atletas de Futsal feminino: descrição e comparação entre as posições TT - Maximum anaerobia power and Futsal female athletes fatigue index: description and comparison between positions. *Rev. bras. ciênc. mov*, *25*(4), 84–91.

Andersson, H. A., Randers, M. B., Heiner-Møller, A., Krustrup, P., & Mohr, M. (2010). Elite female soccer players perform more high-intensity running when playing in international games compared with domestic league games. *Journal of Strength and Condition*, *24*(4), 912–919.

Andersson, H., Raastad, T., Nilsson, J., Paulsen, G., Garthe, I., & Kadi, F. (2008). Neuromuscular fatigue and recovery in elite female soccer: Effects of active recovery. *Medicine and Science in Sports and Exercise*, *40*(2), 372–380.

Barbalho, M. de S. M., Novoa, H. J. D., & Amaral, J. C. (2017). Consumo máximo de oxigênio (vo2) em atletas de futebol profissional de diferentes posições de jogo. *Revista Brasileira de Futsal e Futebol*, *19*(32), 37–41.

Beato, M., Jamil, M., & Devereux, G. (2018). Reliability of internal and external load parameters in recreational football (soccer) for health. *Research in Sports Medicine*, *26*(2), 244–250.

Bourdon, P. C., Cardinale, M., Murray, A., Gastin, P., Kellmann, M., Varley, M. C., Gabbett, T. J., Coutts, A. J., Burgess, D. J., Gregson, W., & Cable, N. T. (2017). Monitoring athlete training loads: Consensus statement. *International Journal of Sports Physiology and Performance*, *12*, 161–170.

Bradley, P. S., Dellal, A., Mohr, M., Castellano, J., & Wilkie, A. (2014). Gender differences in match performance characteristics of soccer players competing in the UEFA Champions League. *Human Movement Science*, *33*(1), 159–171.

Buchheit, M., Manouvrier, C., Cassirame, J., & Morin, J. B. (2015). Monitoring locomotor load in soccer: Is metabolic power, powerful? *International Journal of Sports Medicine*, *36*(14), 1149–1155.

Carling, C., Bloomfield, J., Nelsen, L., & Reilly, T. (2008). The Role of Motion Analysis in Elite Soccer Work Rate Data. *Sports Medicine*, *38*(10), 839–862.

Clemente, F. M., Sarmento, H., Rabbani, A., Van Der Linden, C. M. I. N., Kargarfard, M., & Costa, I. T. (2019). Variations of external load variables between medium- and large-sided soccer games in professional players. *Research in Sports Medicine*, *27*(1), 50–59.

Cohen, A. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2º ed). Routledge.

Cummins, C., Orr, R., O'Connor, H., & West, C. (2013). Global positioning systems (GPS) and microtechnology sensors in team sports: A systematic review. *Sports Medicine*, *43*(10), 1025–1042.

Dalen, T., JØrgen, I., Gertjan, E., Havard, H. G., & Ulrik, W. (2016). Player load, acceleration, and deceleration during forty-five competitive matches of elite soccer. *Journal of Strength and Conditioning Research*, *30*(2), 351–359.

Datson, N., Hulton, A., Andersson, H., Lewis, T., Weston, M., Drust, B., & Gregson, W. (2014). Applied physiology of female soccer: An update. *Sports Medicine*, *44*(9), 1225–1240.

Evangelio, C., Sierra-Díaz, M. J., González-Víllora, S., & Clemente, F. M. (2019). 'Four goals for three players': using 3 vs. 3 small-sided games at school. *Human Movement*, *20*(4), 68–78.

Gaudino, P., Iaia, F., Strudwick, A. J., Hawkins, R. D., Alberti, G., Atkinson, G., & Gregson, W. (2015). Factors influencing perception of effort (session-RPE) during elite soccer training. *International Journal of Sports Physiology and Performance*.

Gauthier-Villars. (1960). *Comptes rendus des séances de la onzième conférence générale des poids et mesures*.

Hennessy, L., & Jeffreys, I. (2018). The current use of GPS, its potential, and limitations in soccer. *Strength and Conditioning Journal*, *40*(3), 83–94.

Hopkins, W. G., Marshall, S. W., Batterham, A. M., & Hanin, J. (2009). Progressive statistics for studies in sports medicine and exercise science. *Medicine and Science in Sports and Exercise*, *41*(1), 3–12.

Maddison, R., & Ni Mhurchu, C. (2009). Global positioning system: A new opportunity in physical activity measurement. *International Journal of Behavioral Nutrition and Physical Activity*, *6*, 73.

Malone, J. J., Lovell, R., Varley, M. C., & Coutts, A. J. (2017). Unpacking the black box: Applications and considerations for using gps devices in sport. *International Journal of Sports Physiology and Performance*, *12*, 18–26.

Marqués-Jiménez, D., Calleja-González, J., Arratibel, I., Delextrat, A., & Terrados, N. (2017). Fatigue and Recovery in Soccer: Evidence and Challenges. *The Open Sports Sciences Journal*, *10*(1), 52–70.

Ohlsson, A., Berg, L., Ljungberg, H., Söderman, K., & Stålnacke, B.-M. (2015). Heart Rate Distribution during Training and a Domestic League Game in Swedish Elite Female Soccer Players. *Annals of Sports Medicine and Research*, *2*(4), 1025.

Park, L. A. F., Scott, D., & Lovell, R. (2019). Velocity zone classification in elite women's

football: where do we draw the lines? *Science and Medicine in Football*, *3*(1), 21–28.

Perrier-Melo, R. J., Barbosa, R. R., Brito-Gomes, J. L. De, Oliveira, S. F. M. de, Fortes, L. D. S., & Costa, M. D. C. (2018). Análise comparativa da potência aeróbia máxima em jogadoras de futebol de acordo com as posições do sistema de jogo. *ConScientiae Saúde*, *17*(4), 463–468.

Rago, V., Brito, J., Figueiredo, P., Costa, J., Barreira, D., Krustrup, P., & Rebelo, A. (2019). Methods to collect and interpret external training load using microtechnology incorporating GPS in professional football: a systematic review. *Research in Sports Medicine*, *00*(00), 1–22.

Ramos, G P, Nakamura, F. Y., Penna, E. M., Wilke, C. F., Pereira, L. A., Loturco, I., Capelli, L., Mahseredjian, F., Silami-Garcia, E., & Coimbra, C. C. (2019). Activity Profiles in U17, U20, and Senior Women's Brazilian National Soccer Teams During International Competitions: Are There Meaningful Differences? *Journal of Strength and Conditioning Research*, *33*(12), 3414–3422.

Ramos, G P, Nakamura, F. Y., Pereira, L. A., Junior, W. B., Mahseredjian, F., Wilke, C. F., Garcia, E. S., & Coimbra, C. C. (2017). Movement Patterns of a U-20 National Women's Soccer Team during Competitive Matches: Influence of Playing Position and Performance in the First Half. *International Journal of Sports Medicine*, *38*(10), 747–754.

Ramos, Guilherme P., Nakamura, F. Y., Penna, E. M., Wilke, C. F., Pereira, L. A., Loturco, I., Capelli, L., Mahseredjian, F., Silami-Garcia, E., & Coimbra, C. C. (2019). Activity Profiles in U17, U20, and Senior Women's Brazilian National Soccer Teams During International Competitions. *Journal of Strength and Conditioning Research*, *33*(12), 3414–3422.

Ramos, Guilherme Passos, Nakamura, F. Y., Pereira, L. A., Junior, W. B., Mahseredjian, F., Wilke, C. F., Garcia, E. S., & Coimbra, C. C. (2017). Movement Patterns of a U-20 National Women's Soccer Team during Competitive Matches: Influence of Playing Position and Performance in the First Half. *International Journal of Sports Medicine*, *38*(10), 747–754.

Reche- Soto, P., Cardona- Nieto, D., Diaz- Suarez, A., Bastida- Castillo, A., Gomez- Carmona, C., Garcia- Rubio, J., Pino- Ortega, J., Reche-Soto, P., Cardona-Nieto, D., Diaz-Suarez, A., Bastida-Castillo, A., Gomez-Carmona, C., Garcia-Rubio, J., & Pino-Ortega, J. (2019). Player load and metabolic power dynamics as load quantifiers in soccer. *Journal of Human Kinetics*, *69*(1), 259–269.

Russell, M., Sparkes, W., Northeast, J., Cook, C. J., Love, T. D., Bracken, R. M., & Kilduff, L. P. (2016). Changes in acceleration and deceleration capacity throughout professional soccer match-play. *Journal of Strength and Conditioning Research*, *30*(10), 2839–2844.

Scott, M., Scott, T., & Kelly, V. (2016). The Validity and Reliability of Global Positioning Systems in Team Sport: A Brief Review. *Journal of Strength and Conditioning Research*, *30*(5), 1470–1490.

Silva, P., Esteves, P., Correia, V., Davids, K., Araújo, D., & Garganta, J. (2015). Effects of manipulations of player numbers vs. Field dimensions on inter-individual coordination during small-sided games in youth football. *International Journal of Performance Analysis in Sport*, *15*(2), 641–659.

Strauss, A, Sparks, M., & Pienaar, C. (2019). The use of GPS analysis to quantify the internal and external match demands of semi-elite level female soccer players during a tournament. *Journal of Sports Science and Medicine*, *18*(1), 73–81.

Strauss, Anita, Sparks, M., & Pienaar, C. (2019). The Use of GPS Analysis to Quantify the Internal and External Match Demands of Semi-Elite Level Female Soccer Players during a

Tournament. *Journal of Sports Science and Medicine*, *18*(1), 73–81.

Teixeira, N. C. G. (2019). Intuição e racionalidade no futebol: os estilos intuitivos e racionais de jogadores de futebol consoante o género e as posições ocupadas em campo. Instituto Universitário do Riod e Janeiro (ISPA).

Trewin, J, Meylan, C., Varley, M. C., & Cronin, J. (2018). The match-to-match variation of match-running in elite female soccer. *Journal of Science and Medicine in Sport*, *21*(2), 196–201.

Trewin, Joshua, Meylan, C., Varley, M. C., & Cronin, J. (2017). The match-to-match variation of match-running in elite female soccer. *Journal of Science and Medicine in Sport*, 6–11.

Varley, M. C., Jaspers, A., Helsen, W. F., & Malone, J. J. (2017). Methodological considerations when quantifying high-intensity efforts in team sport using global positioning system technology. *International Journal of Sports Physiology and Performance*, *12*(8), 1059–1068.

Weston, M. (2018). Training load monitoring in elite English soccer: a comparison of practices and perceptions between coaches and practitioners. *Science and Medicine in Football*, *2*(3), 216–224.

CONCLUSÃO

Existe uma lacuna na literatura cientifica quando falamos sobre futebol feminino, estudos que caminham nexto contexto ainda são incipientes principalmente aqueles que querem determinar o controle das cargas de jogos e treinamento em atletas. Esta lacuna foi apresentada no primeiro estudo que compõe essa dissertação que demostra dos trinta termos mais relevantes em 9,649 estudos sobre analise de desempenho no futebol nenhuma palavra que remetesse ao gênero feminino foi apresentada. Por outro lado, despertou o interesse para uma investigação mais detalhada sobre as demandas físicas que compoõe um jogo de futebol feminino, ao apresentar tópicos relevantes sobre a capacidade física e o uso de tecnologia vestíveis.

Com essa investigação foi possível investigar e responder a pergunta apresentada na descrição da situação problema, sendo: Quais são as variáveis que devem ser consideradas um time de futebol profissional para determinar o controle das cargas?

Pela falta de informações consistentes sobre o controle de carga no futebol feminino, o segundo estudo mostrou que apenas o trabalho de Park et al, (2019) estava preocupado com determinar valores constistentes de velocidade. Em seu estudo Park et al, (2019), demostrou que as zonas de velocidades, aceleração e desaceleração são diferentes do contexto masculino como utilizados em outros estudos como iguais. Portanto, as variáveis que apresentam atividades em alta intensidade durante uma partida parecem ser as mais interessantes para serem exploradas, portando uma aplicação as demandas de volume e intensidade das ações executadas pelas atletas parece ser um caminho animador.

Outro aspecto importante é compreender o comportamento das variáveis durante competições que ocorrem de forma comcomutante no cenário do futebol brasileiro. Desta forma, o estudo 3 caminhou a compreender o comportamento das variáveis relacionadas as demandas de carga externa executadas pelas as jogadores em suas posições no campo de jogo durantes as partidas em duas competições. Assim, foi observado um decréscimo de todas as variáveis de demanda externa coletadas através da tecnologia de GPS em cada posição entre os períodos das partidas, trazendo significância estatística e magnitudes de efeito moderadas que podem estar associadas a um desgaste fisiológico das jogadoras durante as competições.

Estes três estudos apresentados merecem complemento o que poderia trazer variações interessantes para responder a segunda pergunta da descrição problema, Quantas variáveis são necessárias para fazer análises de desempenho de jogadores de futebol? Como sugestão

futura trabalhar para responder essa pergunta pode-se trabalhar com os dados de GPS afim de obter scores, ou seja, variáveis artificiais criadas que podem representar os componentes de força, resistência e cardiovascular inerentes no processo de controle de carga das jogadoras.

REFERÊNCIAS

ALAMAR, B. C. **Sports Analytics: A Guide for Coaches, Managers, and Other Decision Makers**. New York: Columbia University Press, 2013.

APPOLINÁRIO, F. **Metodologia da ciência: filosofia e prática da pesquisa.** 2 ed. ed. São Paulo: Cengage Learning, 2012.

AUGHEY, R. J.; FALLOON, C. Real-time versus post-game GPS data in team sports. **Journal of Science and Medicine in Sport**, v. 13, n. 3, p. 348–349, 2010.

BANGSBO, J.; NORREGAARD, L.; THORSOE, F. The effect of carbohydrate diet on intermittent exercise performance. **International Journal of Sports Medicine**, v. 13, n. 2, p. 152–157, 1992.

BISHOP, D. An Applied Research Model for the Sport Sciences. v. 38, n. 3, p. 253–263, 2008.

BRADLEY, P. S. et al. The effect of playing formation on high-intensity running and technical profiles in English FA Premier League soccer matches. **Journal of sports sciences**, v. 29, n. 8, p. 821–830, 2011.

BRADLEY, P. S. ET AL. High-intensity activity profiles of elite soccer players at different performance levels. **Journal of strength and conditioning research / National Strength & Conditioning Association**, v. 24, n. 9, p. 2343–2351, 2010.

CARLING, C. et al. Comment on "Performance analysis in football: A critical review and implications for future research". **Journal of sports sciences**, v. 31, n. January 2014, p. 37–41, 2013.

CARLING, C.; REILLY, T.; WILLIAMS, A. M. **Performance assessment**. [s.l: s.n.].

CARLING, C.; WILLIAMS, A. M.; REILLY, T. **Handbook of soccer Match Analysis. A systematic approach to improving performance**. [s.l: s.n.].

CARRON, A. V. et al. Cohesion and Performance in Sport: A Meta Analysis. **Journal of Sport and Exercise Psychology**, v. 24, n. 2, p. 168–188, 2002.

CARRON, A. V.; BRAWLEY, L. R.; WIDMEYER, W. N. Measurement of cohesion in sport and exercise. **Advances in sport and exercise psychology measurement**, v. 213–226, 1998.

CARRON, A. V; BRAY, S. R.; EYS, M. A. Team cohesion and team success in sport. **Journal of Sports Sciences**, v. 20, n. 2, p. 119–126, 2002.

CASTELLANO, J.; ALVAREZ-PASTOR, D.; BRADLEY, P. S. Evaluation of research using computerised tracking systems (amisco?? and prozone??) to analyse physical performance in elite soccer: A systematic review. **Sports Medicine**, v. 44, n. 5, p. 701–712, 2014.

D'ORAZIO, T.; LEO, M. A review of vision-based systems for soccer video analysis. **Pattern Recognition**, v. 43, n. 8, p. 2911–2926, 2010.

DAVIDS, K.; ARAUJO, D.; SHUTTLEWORTH, R. Applications of Dynamical Systems Theory to Football. **Science and Football V**, n. i, p. 537–550, 2005.

DAWSON, B. et al. Player movement patterns and game activities in the Australian Football

League. **Journal of Science and Medicine in Sport**, v. 7, n. 3, p. 278–291, 2004a.

DAWSON, B. et al. Comparison of training activities and game demands in the Australian Football League. **Journal of Science and Medicine in Sport**, v. 7, n. 3, p. 292–301, 2004b.

EYS, M. et al. Cohesion and Performance for Female and Male Sport Teams. **The Sport Psychologist**, v. 29, n. 2, p. 97–109, 2015.

FIFA. Financial and Governance Report 2015. n. May, p. 134, 2016.

FIFA, F. I. DE F. A. Women ' s Football. n. May, p. 1–3, 2015.

FUSTER-PARRA, P. et al. A Bayesian network to discover relationships between negative features in sport: A case study of teen players. **Quality and Quantity**, v. 48, n. 3, p. 1473–1491, 2014.

GARGANTA, J. A análise da performance nos jogos desportivos. Revisão acerca da análise do jogo. **Revista Portuguesa de Ciências do Desporto**, v. 1, n. 1, p. 57–64, 2001.

GARGANTA, J. Trends of tactical performance analysis in team sports: bridging the gap between research, training and competition. **Revista Portuguesa de Ciências do Desporto**, v. 9, n. 1, p. 81–89, 2009.

GIL, A. C. **Como Elaborar Projetos de Pesquisa**. 4ª Edição ed. São Paulo: [s.n.].

GUDMUNDSSON, J.; WOLLE, T. Football Analysis Using Spatio-temporal Tools. **Proceedings of the 20th International Conference on Advances in Geographic Information Systems**, p. 566–569, 2012.

HAUGEN, T. A.; TØNNESSEN, E.; SEILER, S. Speed and countermovement-jump characteristics of elite female soccer players, 1995-2010. **International Journal of Sports Physiology and Performance**, v. 7, n. 4, p. 340–349, 2012.

HEWITT, A.; NORTON, K.; LYONS, K. Movement profiles of elite women soccer players during international matches and the effect of opposition's team ranking. **Journal of Sports Sciences**, v. 32, n. 20, p. 1874–1880, 2014.

HUGHES, M.; BARTLETT, R. The use of performance indicators in performance analysis. **Journal of Sports Sciences**, v. 20, n. 10, p. 739–754, 2002.

HUGHES, M.; FRANKS, I. **The Essentials of Performance Analysis An Introduction**. London: Routledge, 2007.

JAMES, N. Notational analysis in soccer: past, present and future. **International journal of performance analysis in sport**, v. 6, n. 2, p. 67–81, 2006.

JONES, N. M. P.; MELLALIEU, S. D.; JAMES, N. Team performance indicators as a function of winning and losing in rugby union. **International Journal of Performance Analysis in Sport**, v. 4, n. 1, p. 61–71, 2004.

JULLIEN, H. Does a short period of lower limb strength training improve performance in field-based tests of running and agility in young professional soccer players? **Journal of strength and conditioning research**, v. 22, n. 2, p. 404–411, 2008.

KRISHNAN, K. **Data Warehousing in the Age of Big Data**. 1st. ed. [s.l.] Morgan Kaufmann, 2013.

LAGO-PEÑAS, C.; LAGO-BALLESTEROS, J.; REY, E. Differences in performance indicators between winning and losing teams in the UEFA Champions League. **Journal of Human Kinetics**, v. 27, n. 1, p. 135–146, 2011.

LIEBERMANN, D. et al. Advances in the application of information technology to sport performance. **Journal of sports sciences**, v. 20, n. 10, p. 755–769, 2002.

MAXCY, J.; DRAYER, J. Sports Analytics: Advancing Decision Making Through Technology and Data. **Fox School of Business**, n. March, p. 27, 2014.

MEMMERT, D.; PERL, J. Human Movement Science Analysis and simulation of creativity learning by means of artificial neural networks. **Human Movement Science**, v. 28, n. 2, p. 263–282, 2009.

MILLER, T. W. **Sports Analytics and Data Science: Winning the Game with Methods and Models**. New Jersey: Perarson Education, 2016.

MOHR, M.; KRUSTRUP, P.; BANGSBO, J. Match performance of high-standard soccer players with special reference to development of fatigue. **Journal of sports sciences**, v. 21, n. 7, p. 519–528, 2003.

MORGULEV, E.; AZAR, O. H.; LIDOR, R. Sports analytics and the big-data era. **International Journal of Data Science and Analytics**, 2018.

O'DONOGHUE, P. Normative Profiles of Sports Performance Peter O'Donoghue School of Sport, Physical Education and Recreation, University of Wales Institute Cardiff, Cyncoed Campus, Cardiff, Wales, CF23 6XD, UK. **International Journal of Performance Analysis in Sport**, v. 5, p. 104–119, 2005.

O'DONOGHUE, P. **RESEARCH METHODS FOR SPORTS**. London: Routledge, 2010.

PRZEDNOWEK, K. et al. Planning Training Loads for the 400 M Hurdles in Three-Month Mesocycles Using Artificial Neural Networks. **Journal of Human Kinetics**, v. 60, n. 1, p. 175–189, 2017.

RAMOS, G. P. PERFIL DA ATLETA DE FUTEBOL DE ELITE NO BRASIL: características antropométricas, físicas e respostas fisiológicas em treinamentos e competições. **Tese de Doutorado apresentada ao Programa de Pós-Graduação em Ciências do Esporte da Escola de Educação Física, Fisioterapia e Terapia Ocupacional da Universidade Federal de Minas Gerais como parte dos requisitos para a obtenção do título de Dou**, 2017.

RAMPININI, E. et al. Validity of simple field tests as indicators of match-related physical performance in top-level professional soccer players. **International Journal of Sports Medicine**, v. 28, n. 3, p. 228–235, 2007.

RAMPININI, E. et al. Technical performance during soccer matches of the Italian Serie A league: Effect of fatigue and competitive level. **Journal of Science and Medicine in Sport**, v. 12, n. 1, p. 227–233, 2009.

ROBERTSON, S.; BACK, N.; BARTLETT, J. D. Explaining match outcome in elite Australian Rules football using team performance indicators. **Journal of Sports Sciences**, v. 34, n. 7, p. 637–644, 2016.

RUSSELL, M.; REES, G.; KINGSLEY, M. I. Technical demands of soccer match play in the English championship. **The Journal of Strength & Conditioning Research**, v. 27, n. 10, p. 2869–2873, 2013.

SILVA, E. J. DE O. **Análise do jogo de futebol: características do processo de transição defesaataque das sequências ofensivas com finalização**. [s.l.] UTAD, 2007.

SMITH, T.; HAMMOND, J.; GILLEARD, W. The use of performance analysis technology to monitor the coaching environment in soccer. **International Journal of Performance Analysis in Sport**, v. 5, n. 3, p. 126–138, 2005.

TRAVASSOS, B. et al. Eco-Dynamics Approach to the study of Team Sports Performance. **The Open Sports Sciences Journal**, v. 3, n. October 2015, p. 56–57, 2010.

TRAVASSOS, B. et al. Performance analysis in team sports: Advances from an ecological dynamics approach. **International Journal of Performance Analysis in Sport**, v. 13, n. 1, p. 83–95, 2013.

WEINECK, J. **Coaching soccer - Conditioning.** Thessaloniki: [s.n.].

WILLIAMS, A. M.; REILLY, & T.; REILLY, T. Talent identification and development in soccer. **Journal of Sports Sciences**, v. 18, n. March 2013, p. 657–667, 2000.

WILSON, D. S.; OSTROM, E.; COX, M. E. Generalizing the core design principles for the efficacy of groups. **Journal of Economic Behavior and Organization**, v. 90, p. S21–S32, 2013.

WRIGHT, C.; ATKINS, S.; JONES, B. An analysis of elite coaches' engagement with performance analysis services (match, notational analysis and technique analysis). **International Journal of Performance Analysis in Sport**, v. 12, n. 2, p. 436–451, 2012.

WRIGHT, C.; CARLING, C.; COLLINS, D. The wider context of performance analysis and it application in the football coaching process. **International Journal of Performance Analysis in Sport**, v. 14, n. 3, p. 709–733, 2014.