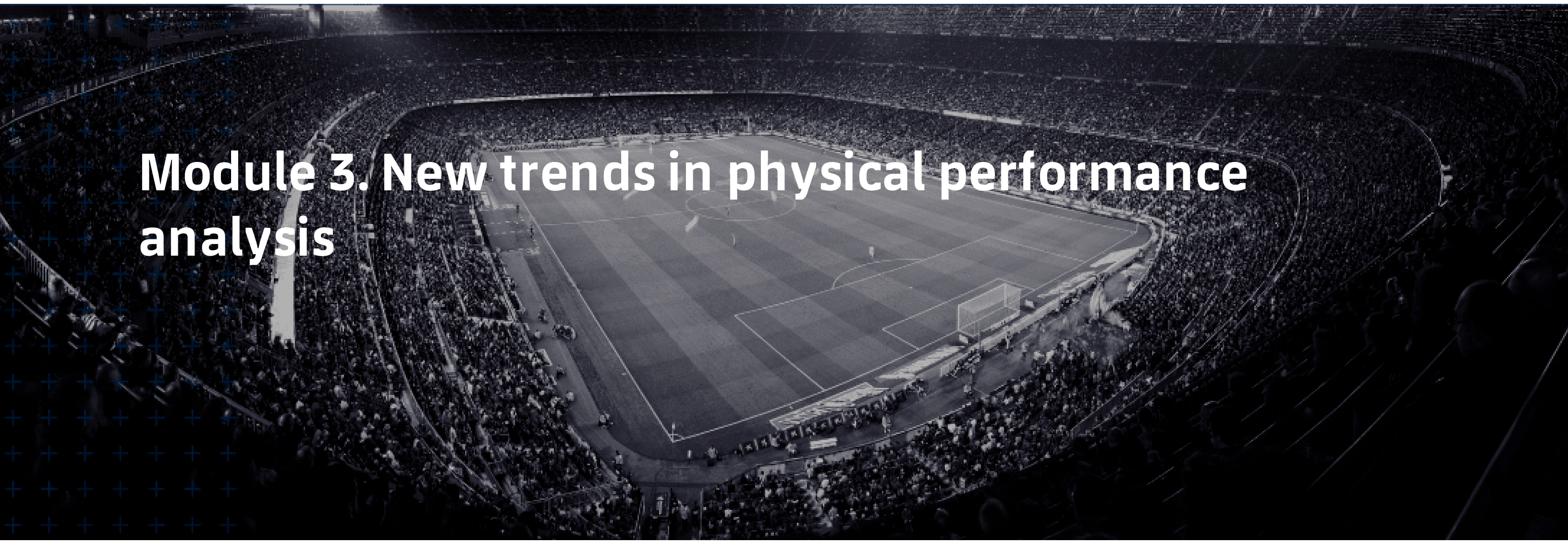


Module 3. New trends in physical performance analysis

BARÇA INNOVATION HUB
Universitas

Unit 3.1 New trends in physical performance analysis

One of the main characteristics that we have highlighted during the course on R and RStudio is their collaborative nature. This aspect is linked to the basic functionalities of R. The software includes a number of functions and add-ons determined in its main installation, but, as we have seen in the reading and video material, other libraries are needed to make our analysis or data processing faster, and our code cleaner or more efficient. These libraries have been developed by other professionals to respond to specific needs. For example, there are specific libraries for the processing and organization of data, which will make it easier to manipulate and change its structure. Other libraries, for example, may focus on time series analysis, the clean up and calculations of time series.

The fundamental principle of these functions is that they are created from other functions and combined in such a way so that we meet our goal as quickly as possible, provided that our data meets the characteristics required by the function.

Because of this, we highlight the great advantage it is to use RStudio. In the same way that professional data analysts, mathematicians and physicists create specific packages to perform analyses in their fields of study, there are multiple sports science professionals who share valuable content with concrete characteristics related to the type and origin of their data, which may facilitate the analysis in our field or with our players.

Thus, when we create functions that can be of help to other professionals and we are willing to share our work with the rest of the community of RStudio users who could benefit from them, their distribution can be done in a relatively simple way.

This module aims, firstly, to highlight a series of methods of physical performance analysis which have been developed in recent years but which have a common characteristic: there are specific packages in RStudio or GitHub that can be downloaded and used as tools in our analyses as sports scientists. It is not our objective to cover each of the new trends in physical performance analysis, but to show that there are tools to apply some of these new trends without the need to buy software or specific technology. Secondly, we aim to encourage you to create new content that could be useful for other professionals and help in the growth and expansion of impact analysis in sports performance.

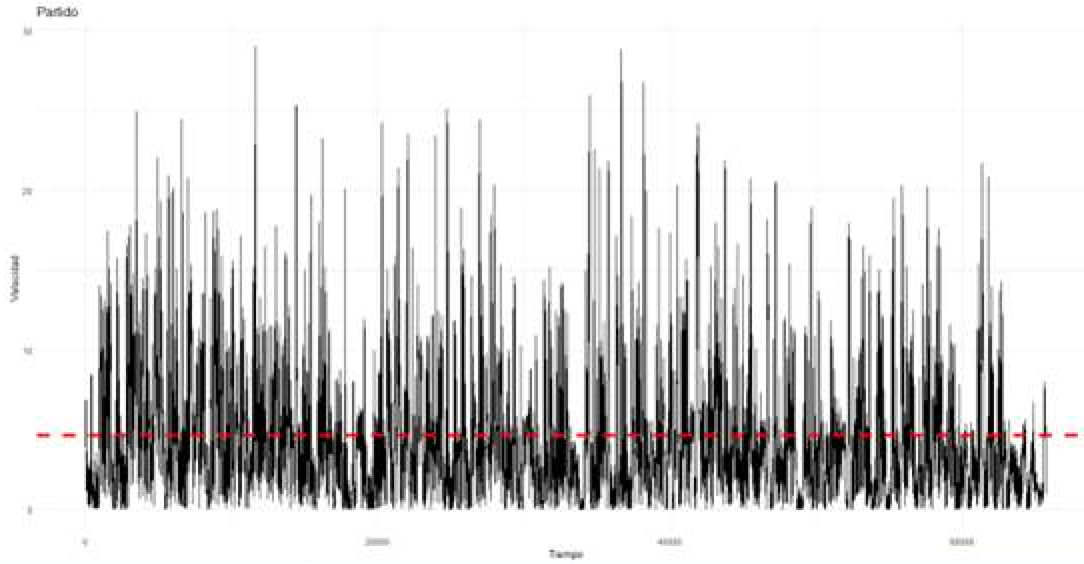
3.1.1 Analysis of periods of maximal and submaximal intensity

Knowledge of the physical demands of sport is the starting point for any methodology on physical development and conditioning. Knowing the conditional stress to which the athlete is subjected during competition and training is key to designing and periodizing their training. On many occasions, competition is the point of reference to determine the players' profile and physical performance.

In intermittent natured sports - in which phases of action and rest are interspersed (Lapuente-Sagarra, 2011), that is, they are not cyclical or of very short duration - the analysis of the average demands of competition does not reflect those phases in which the player is subjected to higher intensities.

For example, in the following graph (Figure 1) we see how, during a match, a football player's speed signal shows that there are phases where the player covers distances at higher speeds (high peaks) which are interspersed by phases of lower activity (lower lines). The red horizontal line shows the average speed of that match (which stands at 4.6 km/h, far from the peaks of more than 25 km/h that the player reaches).

Figure 1: Speed signal recorded by GPS during a football



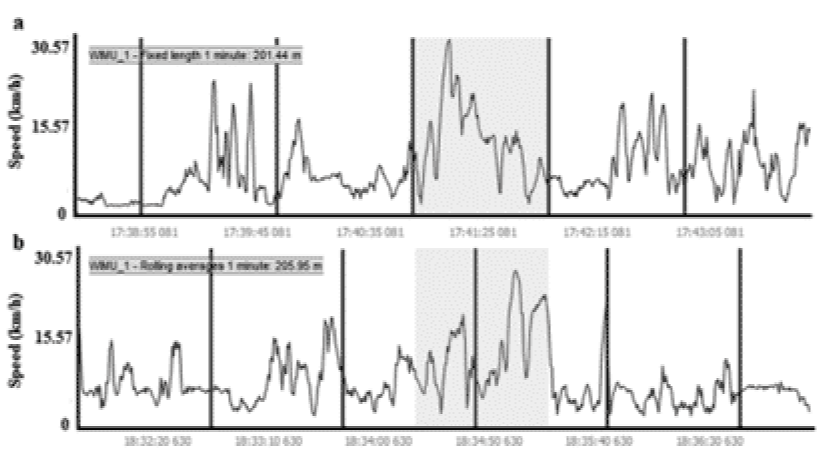
Source: Author's own production

As we can see, the average speed is not representative of the player's physical demands. If we thought so, we would be ignoring that intermittent nature that we highlighted earlier. Therefore, this average value presents a clear limitation if used as a referent of intensity when replicating the demands of competition in our training or introducing similar intensities in rehabilitation or pre-season processes.

To overcome this problem, different approaches have been used to analyse phases of greater intensity than the match average intensity. In these first analyses, it was decided to divide the total duration of the match into fixed time blocks in order to analyse that average speed every 1.3 or 5 minutes. However, Varley et al. (2012) warn that this methodology may not be adequate to detect the most intense phases since these periods can be found between the fixed time blocks.

The method used to detect these periods between fixed time blocks is called the rolling average. Rolling averages allow us to capture the average intensity of the determined duration (1, 3, 5, etc. minutes) throughout the entire time series. For example, if we use a rolling average of 1 minute in a speed signal that gets data every second, at a time point of 1'01", we will have the average intensity from 1" to 1'01"; then, at the time point 1'02", the average intensity we will get ranges from 2" to 1'02", and so on throughout the signal.

Figure 2: Difference between fixed-length (a) and rolling average (b) methods for WCS



Source: Oliva-Lozano et al. (2020), p. 326

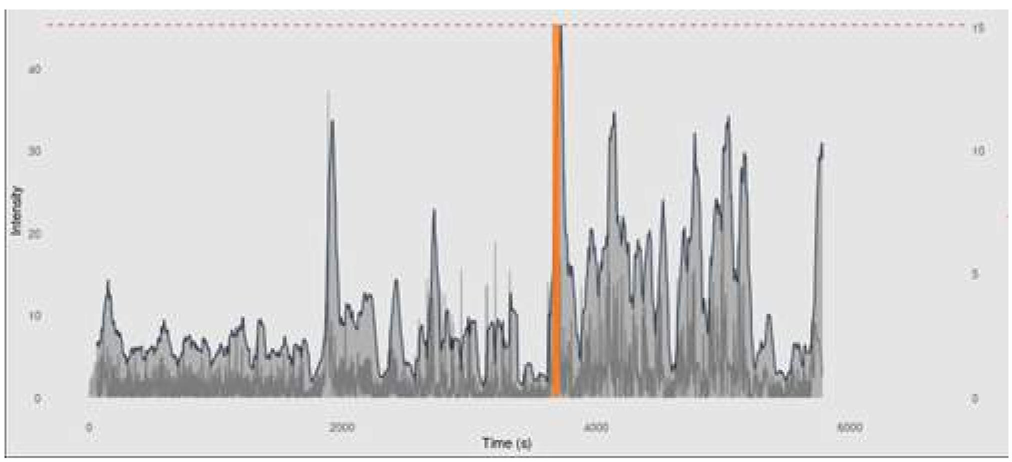
This visualization by Oliva Lozano et al. (2020) perfectly presents these differences. If we want to analyse the minute of greatest intensity of this speed signal, the minute in which the greatest distance has been covered, splitting the intensity signal by fixed windows of 1 minute, the value we get is 201 metres. Looking at the graph, we see that it occurred at minute 4, in the fourth block. However, in the second example, using the rolling average, we get a distance of almost 206 meters, and this is between two blocks of fixed time, which shows the advantages of analysis using rolling averages.

Using this methodology, Garcia et al. (2018) analysed the periods of highest intensity by positional group in football and compared these intensities with those got during training to determine if the intensity stimulus during training was similar to the demands of competition. To do this, they analysed time windows (rolling average durations) of 3, 5 and 10 minutes, for each of the variables.

With these last two examples, we pretend to emphasize the importance of selecting an analysis criteria. First, the intensity we could get would be consistent with the duration analysed (Delaney et al., 2017). And we have to choose it according to a given criteria. In some contexts, the duration of the time window is chosen based on the duration of the tasks we use in training; in other cases, the selection may be justified in relation to physiological adaptations.

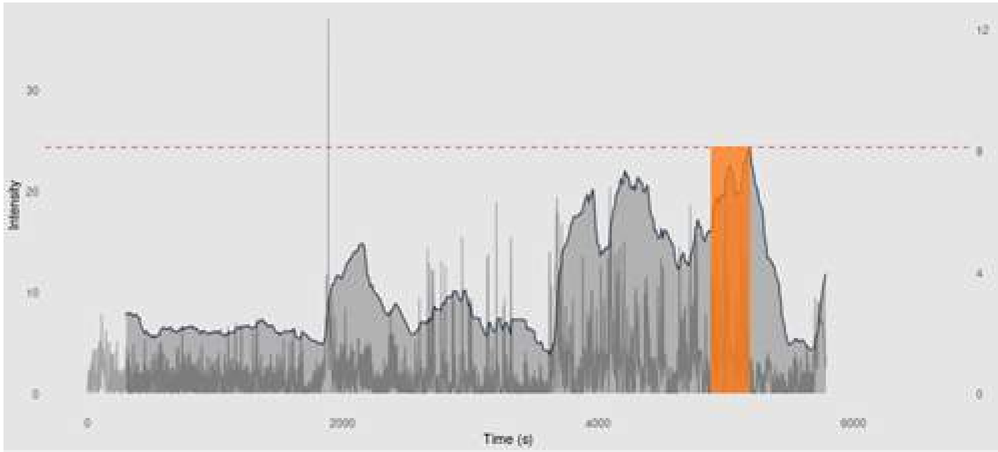
The same intensity signal is shown below using two time windows - in the first graph (Figure 3), it is 1 minute; in the second graph (Figure 4), it is 5 minutes. The intensity signal used in both cases is the speed in km/h; the original speed signal is the one shown in black and blurred in the background; and the rolling average is represented by the blue line and gray shading.

Figure 3: Speed signal using a 1-minute time window



Source: Author's own production

Figure 4: Speed signal using a 5-minute time window



Source: Author's own production

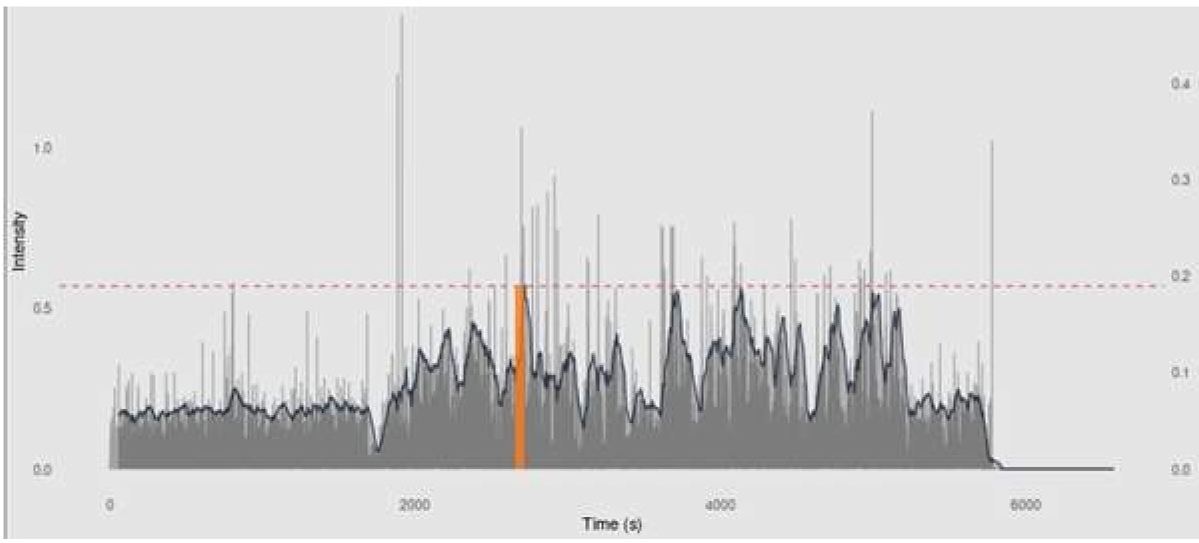
It is clear from the examples that the rolling average makes the signal strength smooth; that is, we go from having a large number of peaks in the signal to a more continuous signal. This is the goal of the rolling average: to summarize a signal with a lot of "noise" (spikes and changes in speed) to give information about what has happened every X amount of time.

The orange shaded bar shows the period of greatest intensity in each of the charts. With this example, we aim to highlight the importance of the criterion to select the time window. The minute of greatest intensity is located at a completely different point in the game than the 5 most intense minutes, and the intensities of each one are also different (represented by the axis on the left).

The second aspect to consider is the variable that will be analysed. The periods of greatest intensity are specific to the variable analysed.

Each of the variables will reflect different aspects of the load (mechanical, locomotor, metabolic), so each of the periods of greatest intensity will be specific to that variable.

Figure 5: Acceleration density signal using a 1-minute time window



Source: Author's own production

This graph portrays the same session, but analysing the acceleration density variable. The time window is 1 minute, the same as in the previous graph, and we can see that this period is in a place in the time series different from the one on the speed signal.

Both points will have direct implications on the results and their interpretation to guide the design or evaluation of tasks that aim to replicate the maximal intensities of the game. The sports scientist’s criteria and context knowledge will be fundamental when deciding the selection criteria.

We can add greater value to these analyses if we follow the work of Ju et al. (2022), in which they recorded the periods of greater intensity and analysed the actions carried out in the match are to discover the type of technical and tactical performance the players and the team have in situations of greater intensity.

However, although the analysis of periods of maximal intensity can be very useful, it also has a series of limitations. Novak et al. (2021) highlighted that there is great variability in the intensities got between matches, and Illa et al. (2020) showed that there are a considerable number of periods in which the player achieves intensities close to the maximal. These two factors show that if only an isolated match value is analysed, other moments in which the player has a high demand without reaching maximal values which could have clear physiological implications may be ignored.

To analyse this type of demand, in recent years the analysis of sub-maximal periods (SubMIP) has been proposed. This analysis aims to analyse all the periods close to a maximal intensity and quantify their intensity, duration and frequency (number of periods).

In the same way that we analyse periods of greater intensity, we have to take into account a series of criteria to analyse sub-maximal periods:

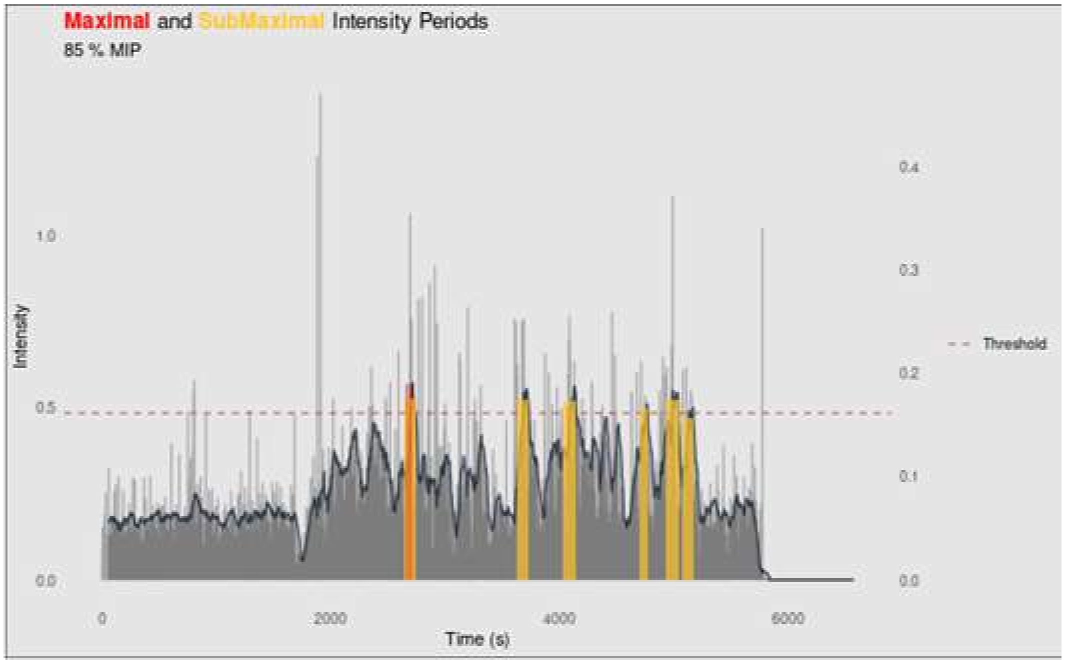
- Duration of the time window: the chosen duration will show the phases close to the maximal of the same duration or a slightly longer one.

- Variable analysed: the phases close to the maximal will be specific to the variable analysed.

- Reference threshold: Here we need to define what we consider to be near-maximal intensity (e.g. 85% of the period of maximal intensity).

- Reference intensity: if we use a time window of 1 minute, the maximal reference intensity should be a time window of 1 minute. We have to choose whether that maximal is from the same session or the same match that we analyse, or an all-time maximal of the position or the player.

Figure 6: Periods of maximal and submaximal intensity



Source: Author's own production

Using the same match as with the acceleration density signal, and also using the 1 minute time window as a criterion and as a period of maximal intensity, the value of the previous graph, and a threshold of 85%, we can see that there are 6 periods in which the player exhibits intensities close to the maximal for periods close to 1 minute.

With this type of analysis, Caro et al. (2022) and Illa et al. (2020) observed differences in the values got across players. They highlighted the importance of analysis for the individualization of the preparation of players; they also compared the values got in the matches with those got in training.

How should we apply these analyses to our data?

One of the strengths of using of RStudio for data analysis is the ability to replicate the methods proposed in different research studies. In the video material, we will see how to replicate this type of analysis from raw GPS data . Moreover, we have highlighted that there are resources at our disposal to facilitate certain processes.

In the case of periods of maximal and submaximal intensity, there is a very simple package, developed as an example, which allows for these analyses to be carried out more effectively, for those professionals who are in a context where they do not have time to elaborate the functions and replicate the steps proposed in academic research; the package is called SubMIP.

We usually download packages or libraries directly from the RStudio interface, but in many cases, we use packages created by individual developers that are not published in CRAN (R Package Library), as they do not meet all their publishing requirements or are still in development. These packages are usually found on GitHub.

GitHub is a platform for developers that allows you to create, store, and share old versions of certain code. We will use a line of code that we will place in our terminal to download any R package we want to use from GitHub.

For maximal and sub-maximal periods, we will use the following code:

```
devtools::install_github('davidpajon/SubMIP')
```

The devtools package must be installed beforehand to allow installations from GitHub. Once installed, the package will meet our goal. It has 3 very simple functions:

- mip() function: it allows you to find the period of maximal strength of the chosen signal and the desired time window. It has two arguments:
  - Metric or variable to be analysed
  - Duration or time window
- submip() function: it allows you to get the number of periods of intensity close to the maximal, their duration and intensity. It has four arguments:
  - Reference intensity value of the period of maximal intensity
  - Duration or time window
  - Metric or variable to be analysed

    - Threshold from which we want to detect periods.

- Submip_plot() function: it uses the same arguments as the previous function and displays a graph of the values obtained as above.

This example aims to show a resource available to the sport scientist that seeks to respond to possible analysis needs based on recent research.

There are two fundamental advantages of GitHub: we can access all the code created by the developer and see how these functions have been created; and, if we need to go deeper into the analysis and we have time, we can use the code base to extend its functionalities and adapt it to the context or the type of data we have. For example, the SubMIP package functions we have described are configured to use raw GPS data at 10 Hz; if our data has a different sample rate, we might adapt the codebase to meet our data requirements. The same would happen for multiple files or if we wanted to analyse multiple variables at once.

Other resources available

As we have highlighted at the beginning of the module, the academic and professional field is constantly developing. New technologies appear and new methods of physical performance analysis, load control, injury risk analysis, among others, are proposed very frequently. RStudio allows us to explore advanced analyses and propose new approaches. Our concerns and the questions we need to answer will influence our needs or search for jobs, code or packages to use. Using code found on GitHub or shared by other professionals is a good practice to get familiar with certain types of analysis and code-writing methods, although we should always know where we have got the codebase from.

Below there are other packages or resources related to the analysis of physical performance, which have been developed by professionals in our field whose objective it is to facilitate the analysis and provide efficient tools to other professionals.

- **Acceleration-speed profile analysis**

  Pearson et al. (2024) have shared a package with several functionalities to analyse the characteristics of the acceleration-speed profile from GPS data or data measured by fraction of time.

  This analysis, originally by Morin et al. (2021), allows for the player's mechanical performance to be measured during the sprint, without the need for tests or specific technology apart from the usual training sessions. The values got through this analysis allow us to determine differences between players, and inter-season and post-injury changes (López-Sagarra et al., 2022).

We can find the information regarding this package at the following link:

GitHub - aaronzpearson/midsprint

- **Dose-response model**

Jovanovic et al. (2020) have shared a package to apply Clarke and Skiba's (2013) analyses, which aim to determine the player's adaptation based on dose control and determined physical test results. This package allows us to appreciate the methodology behind this proposal and use it in our context.

We can find the information regarding this package at the following link:

GitHub - mladenjovanovic/dorem: Dose Response Modeling

- **Biological and chronological age**

Fernandez provides a package in which functions and visualizations can be used to determine the differences between the athlete’s biological and chronological age, among others (R Core Team, n.d.). This is a key concept in talent identification which should be considered by professionals in the physical performance field.

We can find the information regarding this package at the following link:

CRAN - Package matuR (r-project.org)

- **Asymmetries determination**

Pearson (2024), following the work of Bishop et al. (2016), shares a package in which we can apply functions to determine asymmetry in compliance with the authors' requirements.

We can find the information regarding this package at the following link:

GitHub - aaronzpearson/interlimb: An R Package to Assess Interlimb Asymmetries

CONTINUE

Activities

The fundamental principle of software functions is that they are created from other functions and combined in such a way so that we meet our goal as quickly as possible, provided that our data meets the characteristics required by the function.

- True
- False

SUBMIT

In the same way that we analyse periods of greater intensity, we have to take into account a series of criteria to analyse sub-maximal periods. Identify the 4 correct criteria:

- [ ] Duration of the time window
- [ ] Variable analysed
- [ ] Reference threshold
- [ ] Reference intensity
- [ ] Intensity threshold

SUBMIT

What are two of the fundamental advantages of GitHub according to the text?

- [ ] We can access all the code created by the developer.
- [ ] It allows you to share code privately with a select group.

- [ ] It facilitates the evolution and adaptation of the code according to the context or type of data.

- [ ] It offers us an integrated development environment (IDE).

- [ ] It allows for the creation of unlimited public repositories.

SUBMIT

References

Bishop, C., Read, P., Chavda, S. & Turner, A. (2016). Asymmetries of the Lower Limb: The Calculation Conundrum in Strength Training and Conditioning. Strength and Conditioning Journal, 38. https://doi.org/10.1519/SSC.0000000000000264

Caro, E., Vázquez, M., Lapuente-Sagarra, M. & Caparros, T. (2022). Analysis of professional soccer players in competitive match play based on submaximum intensity periods. PeerJ, 10. https://doi.org/10.7717/peerj.13309.

Clarke, D. C. & Skiba, P. F. (2013). Rationale and resources for teaching the mathematical modeling of athletic training and performance. Advances in Physiology Education, 37(2), 134-152. https://doi.org/10.1152/advan.00078.2011.

Delaney, J., Compton, H., Rowell, A., Dascombe, B., Aughey, R. & Duthie, G. (2017). Modelling the decrement in running intensity within professional soccer players. Science and Medicine in Football, 2(2), 86-92. https://doi.org/10.1080/24733938.2017.1383623.

R Core Team. (s. f.). matuR: Athlete Maturation and Biobanding. R Project. https://cran.r-project.org/web/packages/matuR/index.html

Illa, J., Fernández, D., Tarragó, J.R. & Reche, X. (2020). Most Demanding Passages in Elite Futsal: An Isolated or a Repeat Situation? Apunts. Educación Física y Deportes, 142, 80-84. https://doi.org/10.5672/apunts.2014-0983.es.(2020/4).142.10.

Jovanovic, M., Hemingway, B. S. & Swinton, P. (2020). dorem: Dose Response Modeling. R package version 0.0.0.9001. Retrieved from https://dorem.net. doi: 10.5281/zenodo.3757085 Dose Response Modeling • dorem (mladenjovanovic.github.io)

Ju, W., Doran, D., Hawkins, R., Gómez-Díaz, A., Martin-Garcia, A., Ade, J. D., Laws, A., Evans, M. & Bradley, P. S. (2022). Contextualised peak periods of play in English Premier League matches. Biol Sport, 39(4), 973-983. https://doi.org/10.5114/biolsport.2022.112083

Lapuente-Sagarra, M. (2011). Control de la carga semanal de entrenamiento en futbolistas mediante tecnología GPS. Revista de preparación física en el fútbol, (2), 37-43. https://innovafootball.wordpress.com/wp-content/uploads/2013/05/control-carga-semanal.pdf

Morin, J. B., Le Mat, Y., Osgnach, C., Barnabo, A., Pilati, A., Samozino, P. & Prampero, P. (2021). Individual acceleration-speed profile in-situ: A proof of concept in professional football players. Journal of Biomechanics, 123. In press. https://doi.org/10.1016/j.jbiomech.2021.110524.

Novak, A., Impellizzeri, F., Trivedi, A., Coutts, A. & Mccall, A. (2021). Analysis of the worst-case scenarios in an elite football team: Towards a better understanding and application. Journal of Sports Sciences, 39, 1-10. 10.1080/02640414.2021.1902138.

Oliva-Lozano, J., Martín-Fuentes, I., Fortes, V. & Muyor, J. (2020). Differences in worst-case scenarios calculated by fixed length and rolling average methods in professional soccer match-play. Biology of Sport, 38(3), 325-331. https://doi.org/10.5114/biolsport.2021.99706.

Pearson, A. (s. f.). Interlimb. GitHub. https://github.com/aaronzpearson/interlimb

Pearson, A., Chu, D. & Ward, P. (2024). midsprint: midsprint makes building in-game player profiles easy. R package version 0.2.0.

Varley, M., Elias, G. y Aughey, R. (2012). Current Match-Analysis Techniques' Underestimation of Intense Periods of High-Velocity Running. International Journal of Sports Physiology and Performance, 7(2), 183-185. https://doi.org/10.1123/ijspp.7.2.183.

López-Sagarra, A., Baena-Raya, A., Casimiro Artes, M., Granero-Gil, P., & Rodríguez-Pérez, M. A. (2022). Seasonal Changes in the Acceleration–Speed Profile of Elite Soccer Players: A Longitudinal Study. Applied Sciences, 12(24), 12987. https://doi.org/10.3390/app122412987 .