RPE

Friday, April 5, 2019 7:43 PM

DataFest 2019 - Dat…

Data Overview

This document describes how the data were collected and how the files can be used.

Common Characteristics

The data were collected during the 2017-2018 season. There are five files that give different aspects of the games. The data themselves were collected through a variety of means.

- Player level data are provided by the individual athletes themselves and by IMU/GPS devices worn on their vests during games. GPS data may not be available if players are out of range of the satellites. Players are uniquely identified by the PlayerID variable in all data files. Note that while Players 18-21 have not played in any of the games in this dataset, so they can be removed from the analysis completely.
- Data are available on each game played during the season. Games are often organized in tournaments, which consist of up to 6 games. Each game consists of two 7-minute halves (except the final game of a tournament (game 6), which consists of two 10-minute halves). Games can have extra time at the referee's discretion, if play is stopped for some reason during the game. There can be up to three games played on a single day. The order and time of the games is provided.

There were a total of 43 games, and they are identified by the GameID, which indicates the order in which the games were played throughout the season. (GameID=1 is the first game played in the season.)

There are four types of files, described below.

games.csv

Tells you when, where, opponent, and high-level outcomes and events in the game ("box scores").

- **How were data collected:** Information comes from this website: https://en.wikipedia.org/wiki/2017%E2%80%9318_World_Rugby_Sevens_Series
- **How to use:** high-level game information.
- **Links to:** GameID links to gps. Date links to wellness, GPS, Rate of Perceived Effort (RPE).

wellness.csv

Self-reported health and wellness for each player.

- **How were data collected:** self-reported by each athlete. In principle, reported every morning before 8:30am. All values are subjective, but Urine Specific Gravity (USG) is recorded through a sensor. Each athlete may have a different sense of what "typical" means for them, so consider standardizing per athlete.
- **How to use:** provides subjective sense of energy levels. USG can provide evidence of dehydration.
- **Links to:** Date links to games, wellness, RPE, GPS. PlayerID links to RPE, GPS.

rpe.csv

Rate of Perceived Effort. Self-reported workloads for each "session". A session can be a workout (focusing on a particular objective) or a game.

- **How were data collected:** In theory, each player rates herself after each session and/or game. It is easy, however, for players to neglect this when playing back-to-back games. Note that each day there can be multiple "sessions", and that a "session" can be a recovery period, a game, strength & conditioning, etc. There is no way to associate a particular rating with a particular game on days in which multiple games were played.
- **How to use:** Can be used to provide a subjective sense of fatigue. Note that what one player rates "4" for RPE another might rate "7" or any other number, so consider standardizing per player. For many sports analysts, a ratio of acute/chronic training load > 1.2 indicates that the athlete is currently in "high" training load and at an increased risk for injury. A ratio < 0.8 indicates that they are "de-training" or recovering. These are cut-off values based on Australian Football League players.
- **Links to:** Date links to wellness, games, GPS . PlayerID links to wellness and GPS,

gps.csv

Position data for each player during a game.

- **How were data collected:** Data collected from sensors worn by players. Originally, data were collected at 100 Hz (100 times per second), but have been collapsed to 10 Hz. Thus, each second, there are 10 "frames" that provide information on player location and acceleration. Note that we do not know the location of the ball, or the orientation of the playing field. The "z" acceleration is in the up-down direction, x is back-front, y is side-to-side.
- **How to use:** With caution! Note that making plots of location is unlikely to help you understand the role of fatigue unless you first think carefully about aspects of location that might be affected by fatigue. Some large-scale things to consider: can you infer tackles? Coaches usually encourage players to keep space between them.
- **Links to:** Date links to games, substitutions, wellness, RPE. PlayerID links to wellness, RPE. GameID links to games.

Other analytics Agencies

Pull out clear Error ← statsports.com ↓ Catapult

For More Info / Inspiration

⇒ U Sydney technical

Aaron Coutts

W/L

Game win from player standpoint → Y

Engine racing

High speed running 95% — objective

total distance — objective

Load Ratio

look w/ spikes

↳ Correlation w/ injuries