

Wallabies Player Fair Market Value & Wallabies PlayerMarket Value Trend

8th April 2022

Summary

Technical details for the supply of Wallabies PlayerFair Market Value (WPFMV) and Wallabies Player Market Value Trend (WPMVT) indexes.

1 Indexes description

1.1 Overview

Wallabies Player Market Value Trend (WPMVT) is an aggregate of indexes that quantify the daily value change of professional football players through the integration of several factors. The most relevant parameter is the players' performance during the football matches.

The WPMVT's computation can be outlined as follow:

1. Set the target player (professional footballer)
2. Set its starting reference value

1.1.1 Technical WPFMV

1. Transactions preprocessing
 - (a) Filter the non-relevant transaction (value-based)
 - (b) Filter the transaction with missing IDs (player, team, or league ID)
 - (c) Discount the transaction value yearly
 - Examine the average value of transactions year by year
 - Fit with a linear model / consider if normalization of the average value is required
 - Choose the reference year (last year) and discount all the values to the designed year
2. Age discounting: coefficients computation
 - Hypothesis: the age dependency is quadratic, with a parabola having the maximum for an "average" age
 - Establish the coefficients of this parabola with the data
 - Write the formula for changing value due to age:

$$val_{new}(val, age_1, age_2) = val \cdot P(age_1)/P(age_2)$$

where $P(age_x)$ is the parabolic coefficient associate of age factor.

3. Performance discounting
 - Fit linearly the graph " past year performance " vs "transaction"

- Write the formula for value change based on the performance:

$$val_{new}(val, per_1, per_2) = val \cdot C \cdot per_1 / per_2$$

Where C is the fit coefficient for the performance factor.

4. Discounting based on starting and arrival league.

- Cluster all the transfers for the starting/arrival league
- Examine the league average value differences
 - Take all the relevant transactions within each market
 - Calculate the average
 - Take the coefficient from league A to league B
 - Compute the ratio between the average value of transfers in the two different leagues;
- Application of a coefficient for the league where the football player plays;

1.1.2 WPFMV

The WPFMV is an aggregate of non-significant benchmarks created to estimate the value of a professional football player in a Fair Market environment.

Given a target professional football player, to determine the WPFMV we proceed with the following algorithm:

1. We identify the best comparable (football players) associated with the target player through the algorithm *Wallabies Player Comparability*
2. We determine which of these players have been transferred in the market during the last 5 years; for each of these:
 - (a) we discount the value from the time in which transfer to the current season occurred, taking into consideration:
 - Trend of the player transfer market between reference seasons
 - Age of the player between reference seasons
 - Trend of the player's performance between reference seasons
 - (b) we discount the value of the comparable player to the target player taking into consideration:
 - 2 players' age of 2 players
 - 2 players' performance
 - Players' starting league
 - Starting team's performance of the two players
3. We compute a weighted average of the previous values. The weights are obtained from the distance of the players according to the algorithm *Wallabies Player Comparability*.

1.1.3 WPMVTI

The WPMVTI (Wallabies Player Market Value Trend Index) is an aggregate of non-significant benchmarks, created to estimate the change in the market value of a professional football player during the season.

Starting from a defined market value for a footballer, provided by the WPFMV index, the WPMVTI index computes the discounted value resulting from the variation of the value after each match. This change, added to the pre-match valuation, gives the new index value.

Finally, the evaluation of the market value will be computed through a moving average process, with decreasing exponential weights during the time, of the last prediction and the previous two.

The aim of the moving average process adoption for the evaluation is to soften fluctuations, keeping the same growth/decrease trend of the prediction.

The drivers of WPMVTI index variation are the following:

- Age
- Role of the player
- The actual target player valuation
- Performance

A player does not receive the performance rating in a match, if one of these 3 conditions happens:

- The player is not convocated
- The player does not play
- The players play under 12 minutes

If the player does not receive a performance rating for a specific match, the market value remains unchanged.

1.2 Market Description

1.2.1 WPFMV

The aggregate of indexes WPFMV aim is to evaluate the value of a player in a Fair Market condition.

To perform this activity, the WPFMV is based on the history of the occurred market transactions, assuming that every single transaction cannot be considered "fair", taking into account the total of the transactions, we obtain the best possible approximation of a fair market system.

1.2.2 WPMVTI

The indexes family WPMV TI describes the change in the market value of a target player after a match.

The variation is due to factors related to individual performance, and general features of the player.

This model can be applied to any player where is possible to obtain performance value age and role for each game played.

The model was "trained" with data for the 2020/21 season on players between the ages of 18 and 33 and an initial market value (provided by the WPFMV index) above 5 million euros.

In addition, the same players, during the season 2020/21, must have played in one of these 5 leagues:

- Italy - Serie A
- Spain - Primera Division
- England - Premier League
- Germany – Bundesliga
- France - Ligue 1

Wallabies Srl has the data of 43 national leagues and the two main European competitions, with the corresponding difficulty coefficients assigned to each league. This allows replicating the WPMVTI index on all players who play in these leagues and of which it is possible to calculate a performance value for each match.

2 Data and methodology

2.1 Data

2.1.1 WPFMV

The data used by the model to calculate the WPFMV index, are the history of player sales transactions.

This information is collected manually and by the crossover of the information from different sources such as:

- Specialized sites
- Specialized media
- Company press releases

The information extracted are as follows:

- Transacted player
- Age at time of transaction
- Role
- Team and starting league
- Arriving team and league
- Value of the transaction

To these are added other information related to the target player extrapolated internally as:

- History of performance
- History of minutes played per game
- History of games played per season
- History of the team's average position in the league

These data are used as variables in the computation of the Fair Value (definition of the WPFMV index), through the discounting processes of transactions in the database.

2.1.3 WPMVTI

The data used by the model for the calculation of the index WPMV TI, are obtained from two sources:

- InStat. Information related to age, role, and performance are extrapolated directly from InStat.
- Index WPFMV. The value of the index WPMV TI is initialized via the index WPFMV.

We analyze each variable in input to the model individually.

Age

The age variable is obtained by the date of birth, contained in the birth date variable. This information is provided by InStat. Age is expressed in years and refers to the years entirely turned.

Example: On "12 - 01 - 2022", a player born on "15 - 12 - 1995" is 26 years old.

Market Value

In the calculation of the player's market value change, after a match is played, it is also used the data related to the current rating related to the market value that the player had before the match played.

This variable assumes the last predicted value via the WPMV TI index. Since making the first prediction of the WPMV index TI is not possible, the first reference value used is the one deriving from the valuation of the WPFMV index.

Performance

The *Performance* is the quantitative evaluation of how a player played in a game. The Performance is computed through the Wallabies Ranking Model, a regression model, which receives as input a series of surveys and statistical data related to a player in a given match and it returns a numerical value between 0 and 10.

In order to standardize the different performances coming from different championships, these values are multiplied by the coefficient of difficulty of each competition of which they are part.

This process implies the standardization of the performance related to the difficulty of the championship.

If a player does not take part in a match or plays less than 12 minutes, the performance for that match will not be assigned.

Position

The Position variable is extracted directly from the provider's data through a process that, given a match, calculates the number of minutes the footballer plays in each position. The assigned position is the one in which the footballer plays most of the minutes.

Using the provider's data, it is possible to get the minutes played in the following positions:

- Goalkeeper
- Defender
- Wing-Back
- Midfielder
- Inside-Midfielder
- Wide-Midfielder
- Attacking-Midfielder
- Wing-Forward
- Forward

Depending on the minutes played in one position, the relative position will be assigned.

In case a player has played the same minutes in two different positions, will be assigned the first one that appears as in the list above.

Example: if a player, in a game, plays 45 minutes as Wing Back and 45 minutes as a Wide - Midfielder, will be assigned the position of Wing Back because it is the first of the two that appears in the list, starting from the top and scrolling down.

The Goalkeeper role will not appear since it was decided to not produce any WPMV TI index for Goalkeepers. In the unlikely situation where one of the players on whom the index was built, had to play on goal, the last role he played in the previous game is confirmed.

2.2 Methodology

2.2.1 WPFMV

2.2.2 WPMVTI

The methodology for the WPMV TI index is composed of several steps:

- Collection of data
- Transformation of data
- Choice of input features
- Training phase of the model
- Phase of Prediction

The first phase is related to the design and training of a model for the WPMV TI index. The second phase consists in predicting the change in the WPMV TI index and, knowing the initial value of the index, it is possible to provide the discounting of the index after each match played.

Data collection

The first operation is Data Collection. At the end of the match, we are able to derive the data that will be used in the WPMV TI index discounting model.

- **Position.** From the provider's data related to the match, the role computation process is performed through an internal code that assigns the role based on the minutes played (Goalkeeper, Defender, Wing-Back, Midfielder, Inside-Midfielder, Wide-Midfielder, Attacking-Midfielder, Wing-Forward, and Forward). Next, as explained in paragraph 2.1.2, under the heading Position, will be assigned the role used in the model.
- **Age.** From the provider's data related to the date of birth for every single player, we calculate the age on the day on which the match occurred.
- **Value.** This variable is calculated by taking the last available value of the WPMV TI index (the value of the index before the update).
- **Performance.** Using the data provided by InStat, the performance is calculated via the Wallabies Ranking Model, weighing it by the difficulty coefficient of the league.

Data Transformation

The second operation is the Transformation and the Normalization of the previous point to obtain those values that will be used by the model.

- Creation of the feature *ln Value*. This is achieved by taking the logarithm on basis e of the feature value

$$\ln Value = \log_e(value)$$

- Creation feature *ln Value*². This is achieved by taking the squared of the feature *ln Value*

$$\ln value^2 = (\ln Value)^2$$

- The variable position is used for the choice of algorithm. Depending on the position assigned to a player, his data is normalized according to different parameters.
- Creation of the feature value performance. This is obtained by calculating the difference between the normalized variables of performance and age. Subsequently, we obtain the normalization of this.

Choice of input features

After the work of data collection and creation of supporting variables, we select what will be the variables used for the computation of the WPMVTI index:

- *age*
- *ln Value*²
- *Performance*
- *Value performance*

The training phase of the model

For each of the above roles, an ad-hoc model is trained as follows:

- We filter players who in the 2020/21 season, in the top 5 leagues, have played in the role of reference.
- We calculate their WPFMV index at the beginning and the end of the season
- We train a regression model that determines the price change after each individual performance.
- This model is optimized by trying to minimize the MSE (Mean-Square-Error) between the end of season value calculated in this way and that given by the WPFMV index, on the entire sample of players.

As a result of this operation, the coefficients are determined for each variable that will be later used in the prediction phase.

Prediction Phase

The initial value of each WPMV TI index is equal to the value of the WPFMV index. Then, after each match, the value of the WPMV TI index is discounted using an AI algorithm.

Performed these operations of creation of new variables, we select and carry out the operations of standardization, taking into consideration the values contained in Tab 2.2, for the following features:

- *age*
- *ln Value²*
- *Performance*
- *Value performance*

The third step is to calculate the prediction of the model. Each role is matched to a differential algorithm with its coefficients combined with the input variables

Calculated this value named Prediction, we will calculate a new value computed by adding Prediction and ln value, and applying an exponential transformation:

$$Prediction2 = e^{(\ln value + Prediction)}$$

Finally, the discounting of the WPMV TI index is obtained by applying a moving average system with exponential weights between the predicted value (Prediction2) and the previous value of the index:

$$WPMV TI_t = \frac{2}{3} \cdot Prediction2 + \frac{1}{3} \cdot WPMV TI_{t-1}$$

Where:

- $WPMV TI_t$ is the new update value of the index $WPMV TI$.
- Prediction 2 is the value obtained as described above.
- $WPMV TI_{t-1}$ is the index value $WPMV TI$ before the game t .

3 TICKER

3.1 Nicolò Barella

- Ticker: NICBAR01
- Link: <https://prod.wallabies.it/market-value/NICBAR01>

3.2 Erling Braut Haaland

- Ticker: ERLHAA02
- Link: <https://prod.wallabies.it/market-value/ERLHAA02>