

WHO ARE THE REAL TOP SCORERS: ALGORITHMICALLY RESORTING A TOP SCORER LIST BY A LIMITED SET OF PREFERENCE ASSUMPTIONS

KJETIL K. HAUGEN

Abstract. This article presents an algorithm, which by taking both goals scored as well as matches played into account, is meant to provide “more sensible” top scorer lists in football. The algorithm computes Upper Hulls recursively in order to provide a new and improved list. The complexity of the algorithm is reasonable, and should imply practically feasible execution on a modern computer. The proposed algorithm may also be an interesting alternative to academic scholar ranking. It has interesting properties for instance compared to the h -index.

1. INTRODUCTION

It does not come as a big surprise that various types of rankings are popular in sports, [16]. The output from sport activities is normally highly quantitatively defined, and the possibility of sorting and making rankings of the “best” and “worst” is evident. It is perhaps likewise unsurprising that football, being the clearly most popular [7, 9], of all professional sports, contains many rankings.

Among the most prestigious ranks in football, are the top scorer ranks. Such ranks exist in local national leagues, as well as in international leagues and tournaments. Habitating the top of such lists signals glory and in most situations a future with limited financial challenges. See for instance [15] for a typical example of such lists.

An interesting feature of such top scorer lists is that there is at least one more dimension which is reasonable to take into account in the ranking – the number of games played by the players. Many, see for instance [17] and [18], argue that the best scorers are those who are able to achieve the maximal amount of goals per match they play. However, such an idea may provide weird rankings. After all, a football player scoring 3 goals in his national debut which turns out to be his only national match, would typically top such a list in his country. One could perhaps try to add constraints of at least a minimal number of games played, but such a ranking strategy would inevitably lead to an “advantage” for the players playing the fewest matches.

In the practical world of football, playing many matches is a positive signal for all players, also goal scorers. As such, it may be more sensible to view this

MSC(2010): primary 52B55, 11H06; secondary 62F07.

Keywords: top scorer ranking, convex combination, convex hull, upper hull.

Comments and suggestions from the anonymous referee are gratefully acknowledged.

problem as the following problem. Assume you measure quality in more than one dimension. Would it still be possible to provide a one-dimensional ranking? Given so, what is a sensible way of doing it? This problem is addressed by Wesson in [14], who suggests to rank by the function $f(g, m) = g^2/m$, where g is the number of goals scored by a given player, and m is the number of matches played by the same player.

Academia seems to have found her solution, the Hirsch [8] (or h)-index. Here, both the number of published articles as well as “research market popularity”, typically measured by citations by other authors, seem to be reasonable dimensions to maximise. The so-called h -index attempts to use a single ranking to embrace both dimensions – article and citation counts. For being a quite simple and somewhat limited measure, the h -index has proven to be very popular among scholars.

The main point in this article, is to suggest an alternative way of ranking top scorers (and maybe also scholars), different from both Wesson’s and Hirsch’s proposed ranking methods. In the next section, a case from Norway is used to illustrate the proposed method (Section 2), while the underlying mathematics is discussed in Section 3. Section 4 concludes, discussing algorithm complexity, algorithm characteristics, academic scholar ranking as an alternative use and practical feasibility.

2. A NORWEGIAN TOP SCORER CASE

The idea underlying the proposed method is in some ways inspired by financial theory methods like Time [5] or Stochastic Dominance [6] or Markowitz Mean-variance Portfolio Theory [11]¹. The idea is to make a limited set of “reasonable” assumptions on how a top scorer could be characterized, and then apply these assumptions in order to resort a typical top scorer list. To illustrate the method, the all time Norwegian top level top scorer list is used. This list is given in Table 1.

As Table 1 indicates, the list is sorted by goals only, which is the normal choice by most publishers of top scorer lists. Table 1 also contains a column for the number of matches, which will turn out to be an essential added information in forthcoming paragraphs.

The main assumption can be established as:

Assumption 1. \mathcal{A} : For any two players i, j : if $m_i = m_j$ and $g_i > g_j$, then player i is a better goal scorer than j .

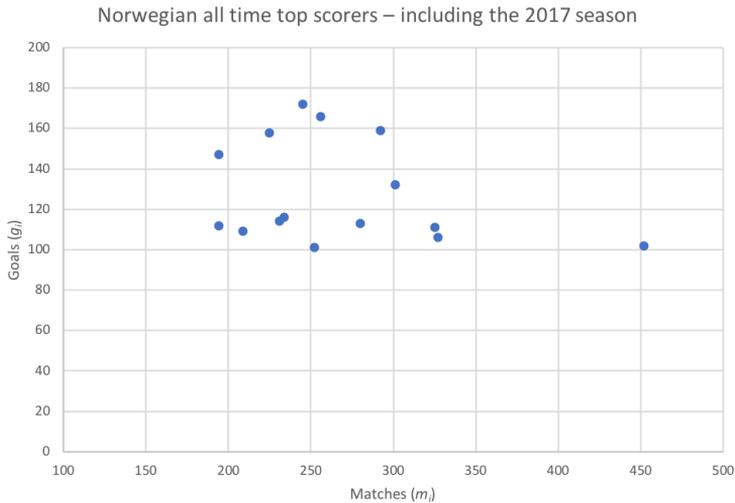
In Assumption 1, g_i is the number of goals scored by player i and m_i is the number of matches played by player i for all players $\in \{1, 2 \dots n\}$. Most people with a reasonable interest in football would accept such an assumption as sensible. In straightforward terms, it only states that if two football players have played the same number of matches, the one with the highest aggregated goal score should be considered a better scorer.

To continue, it turns out to be interesting to investigate a plot of the relevant data. Figure 1 contains an XY -diagram of the two variables (g_i, m_i) .

¹Some might also find similarities to Multi-criteria Decision Making Methods, see for instance [13].

Table 1. All time Norwegian top scorer list (15 first, sorted by number of goals), updated to the end of the 2017 season.

Rank	Name	Goals	Matches
1.	S. Rushfeldt	172	245
2.	H. M. Brattbakk	166	256
3.	P. Belsvik	159	292
4.	O. Iversen	158	225
5.	P. Kristoffersen	147	194
6.	F. Johnsen	132	301
7.	T. Helstad	116	234
8.	A. Pedersen	114	231
9.	B. Sæternes	113	280
10.	J. Flo	112	194
11.	A. Sundgot	111	325
12.	J. Fuglseth	109	209
13.	S. Mathisen	106	327
14.	M. Berre	102	452
15.	O. M. Årst	101	252

**Figure 1.** Matches plotted against goals for all players on the list.

Observing Figure 1, it is relatively easy to see that only two points make Assumption 1 applicable. The points (or players) ranked 5 and 10 in Table 1 are the only players who have played an equal number of matches – 194. Unfortunately (and by all means very logically) these two players are already ranked as they should, so Assumption 1 does not induce any re-ranking. In order to progress, a new (corollary) assumption must be made:

Assumption 2. \mathcal{A}' : Convex combinations of pairs of players are allowed.

Obviously, convex combinations of players' quality in the two given dimensions are a theoretical construct. However, a ranking is a theoretical construct in itself. Opening up for using combined qualities of more than one player when the task is to rank many players seems like a reasonable assumption to make. The first order consequence of Assumption 2, \mathcal{A}' , is illustrated in Figure 2²:

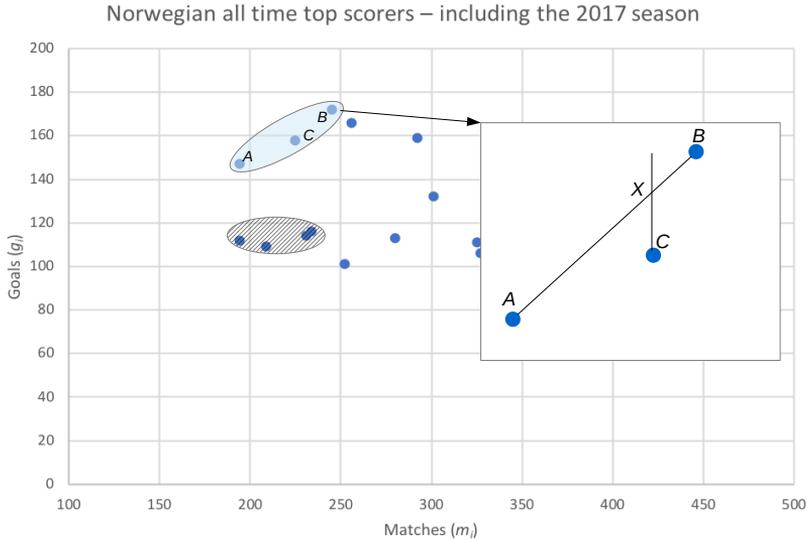


Figure 2. The consequences of allowing convex combinations of players.

Figure 2 equals Figure 1 with the exception that three data points are identified and labelled **A**, **B** and **C**. Furthermore, the convex combination between **A** and **B** is formed as well as a point **X** found by the intersection of the convex combination between **A**, **B** and a normal (relative to the g -axis) from point **C**. Necessary data to calculate point **X** is found by the relevant player data in Table 1 as shown in Table 2 Now, by elementary high school algebra, point **X** is easily calculated.

Table 2. Data for points A, B and C in Figure 2.

Rank	Name	Goals (g)	Matches (m)	Fig. 2
1.	S. Rushfeldt	172	245	A
4.	O. Iversen	158	225	C
5.	P. Kristoffersen	147	194	B

First, the straight line through **A** and **B** is found as (rounding of results is used):

$$g = \left(\frac{147 - 172}{194 - 245} \right) \cdot m + 172 - \left(\frac{147 - 172}{194 - 245} \right) \cdot 245 \Rightarrow g = 0.49m + 51.90.$$

²Note that the data in Figure 2 are the same as those in Figure 1. Some are however hidden due to a sub graph zoom window.

As the point \mathbf{X} is constructed such that it has the same m -coordinate (number of matches) as player \mathbf{C} , it is easy to find it as: $(m^{\mathbf{X}}, g^{\mathbf{X}}) = (225, 0.49 \cdot 225 + 51.90) = (225, 162)$.

Now, the beauty of the argument is evident. Player \mathbf{C} is suddenly “dominated” by the combined \mathbf{A} and \mathbf{B} players. This combination has scored more goals for the same number of matches – Assumption 1 is applicable.

Given the elimination of point \mathbf{C} , more points can be eliminated. In fact all points geometrically placed below the line joining \mathbf{A} and \mathbf{B} should be ranked under players \mathbf{A} and \mathbf{B} in the interval $[m_{\mathbf{A}}, m_{\mathbf{B}}]$. This is shown by the four points within the hatched ellipse in Figure 2.

There is of course no reason to stop here, more potentially “dominating” convex combinations could be constructed, and in order to rule out the maximal amount of players, a situation like that in Figure 3 seems obvious.



Figure 3. Convex combinations of several pairs of players.

In Figure 3, all points **except** \mathbf{A} , \mathbf{B} , \mathbf{D} and \mathbf{E} are dominated by pairwise convex combinations of \mathbf{A} , \mathbf{B} , \mathbf{D} and \mathbf{E} .

This means, that the four players \mathbf{A} (P. Kristoffersen), \mathbf{B} (S. Rushfeldt), \mathbf{D} (P. Belsvik) and \mathbf{E} (M. Berre) should be ranked in front of all other players. Hence, they can be removed from the data set and constitute the first four re-ranked players, as indicated in Table 3. Obviously, our method does not provide information about any re-ranking between these four players, so the original rank, based on goal score is kept.

At this point, 4 players from the original ranking are re-ranked in front of all players. Then, it is straightforward to realize that we can repeat the above process, simply by removing these 4 players and perform a similar forming of pairwise convex combinations on the remaining 11 players. Without going through all tedious details, the results of this process are summed up in Figure 4.

Table 3. The first step of a re-ranking algorithm (R. means Rank).

New R.	Orig. R.	Name	Goals (g)	Matches (m)	Fig. 2
1.	1.	S. Rushfeldt	172	245	B
2.	3.	P. Belsvik	159	292	D
3.	5.	P. Kristoffersen	147	194	A
4.	14.	M. Berre	102	452	E

Step 2 in Figure 4 involves removing and re-ranking 6 players – see Table 4, while Step 3 involves removing and re-ranking 3 players – see Table 5.

Table 4. The second step of a re-ranking algorithm.

New Rank	Orig. Rank	Name	Goals (g)	Matches (m)
5.	2.	H. M. Brattbakk	166	256
6.	4.	O. Iversen	158	225
7.	6.	F. Johnsen	132	301
8.	10.	J. Flo	112	194
9.	11.	A. Sundgot	111	325
10.	12.	S. Mathisen	106	327

Table 5. The third step of a re-ranking algorithm.

New Rank	Orig. Rank	Name	Goals (g)	Matches (m)
11.	7.	T. Helstad	116	234
12.	9.	B. Sæternes	113	280
13.	12.	J. Fuglseth	109	209

At this point (refer to Figure 4, Step 3), only two players are left, and the algorithm stops. These two players; A. Pedersen and O. M Årst will then end up as number 14 and 15 respectively in the final sort. Table 6 sums up:

A quick look at Table 6 shows that the proposed re-rank has made a significant difference. For instance, originally placed as number 14, Morten Berre, is now moved 10 places up to the 4th position. On the other hand, A. Pedersen, originally placed 8th, has moved 6 places down to 14. Only two players are not re-ranked, S. Rushfeldt originally the best, and still the best after the re-rank, and O. M. Årst originally last and still last after the re-rank. The fact that the rank of these players are unchanged is a consequence of the method. See Section 4 for further discussions on this fact.

3. THE UNDERLYING MATHEMATICS AND A FORMALIZED RANKING ALGORITHM

In this section, a more formalized mathematical foundation for the ranking method described in Section 2 is presented.



(a) Step 2



(b) Step 3

Figure 4. The two final steps of a re-ranking algorithm.

The Convex Hull of a finite point set S can be defined (see for instance [10]) as:

$$\text{Conv}(S) = \left\{ \sum_{i=1}^{|S|} \alpha_i \mathbf{X}_i \mid (\forall i : \alpha_i \geq 0) \wedge \sum_{i=1}^{|S|} \alpha_i = 1 \wedge \mathbf{X}_i \in S \right\}.$$

Table 6. The final re-ranked list of Norwegian top scorers.

New Rank	Orig. Rank	Name	Goals (g)	Matches (m)
1.	1.	S. Rushfeldt	172	245
2.	3.	P. Belsvik	159	292
3.	5.	P. Kristoffersen	147	194
4.	14.	M. Berre	102	452
5.	2.	H. M. Brattbakk	166	256
6.	4.	O. Iversen	158	225
7.	6.	F. Johnsen	132	301
8.	10.	J. Flo	112	194
9.	11.	A. Sundgot	111	325
10.	12.	S. Mathisen	106	327
11.	7.	T. Helstad	116	234
12.	9.	B. Sæternes	113	280
13.	12.	J. Fuglseth	109	209
14.	8.	A. Pedersen	114	231
15.	15.	O. M. Årst	101	252

If S is taken to be the points in the Norwegian top scorer list from the previous section, $\text{Conv}(S)$ will be the polyhedral set defined by the red and blue **boundary** lines in Figure 5.

**Figure 5.** The Convex Hull of the Norwegian top scorer list.

Note that the $\text{Conv}(S)$ **boundary** is partitioned in two parts in Figure 5, a red part on top, and a blue part on the bottom. These two parts are defined as the Upper Hull, and the Lower Hull respectively – see for instance [4]. Using a typical

definition, for instance as in [4], the Upper Hull can (given a two dimensional S) be defined as:

$\text{UH}(S) =$ The part of *theConv*(S) **boundary** which is visible from above.

Comparing the Upper Hull in Figure 5 with the case studied in Section 2, it ought to be obvious that each step of the algorithmic skeleton implies finding $\text{UH}(S)$.

As a consequence, a formalized algorithm of the proposed re-ranking can be defined:

Algorithm 1: An algorithm for re-ranking a typical top scorer's list in football

Input : The set S of a top scorer list containing points
 $(g_i, m_i), i \in \{1, \dots, n\}$

- 1 Find $S' = \text{UH}(S)$, by equation (3.1): as a subset of indices, $S' \subseteq S$, and place S' consecutively into the re-ranked top scorer list
- 2 $S := S \setminus S'$
- 3 **if** $S = \emptyset$ **then**
- 4 | **GOTO** 7;
- 5 **else**
- 6 | **GOTO** 1;
- 7 **end**

Although the algorithm above should be self explanatory, a quick explanation may be helpful. At **line 1**, the first Upper Hull is found as a set of indices who are a subset of the original list $\{1, \dots, n\}$. Referring back to the example in Section 2, $S' = \{1, 3, 5, 14\}$. S' is then placed consecutively into the re-ranked list such that players originally ranked 1,3,5 and 14 now are ranked 1,2,3 and 4. Then, at **line 2**, the new (remaining player) set S is formed by applying a set-minus operation; $S := S \setminus S'$. That is, $\{1, 2, \dots, 15\} \setminus \{1, 3, 5, 14\} = \{2, 4, 6, 7, \dots, 13, 15\}$. The algorithm repeats itself until it stops. This happens one step after the end of the example in Section 2. Consider Figure 4, (b) Step 3. After this Upper Hull is identified, only two points remain in S . When this new S is input to a new $\text{UH}(S)$ calculation, $S' = S$, and $S \setminus S' = \emptyset$, and the algorithm stops. The final output of the algorithm, the re-ranked list, will now contain the vector equal to the second column (**Orig. Rank**) in Table 6.

4. DISCUSSION AND CONCLUSIONS

4.1. Complexity of the algorithm

For any algorithm, some notions on complexity is important to judge. The main workload of the algorithm presented in Section 3 is related to the Upper Hull computations. Each pass of the algorithm contains the identification of an Upper Hull, and obviously, the complexity of these operations will contain the most significant computational workload.

Fortunately, research in computational geometry already provides many algorithms for efficient computation of Convex as well as Upper Hulls. In addition, complexity of these algorithms is already well studied. See for instance [2,3]. Some of these algorithms, for instance [1] has the nice property that they find the Upper Hull first, which of course is especially convenient in this case, where the Upper Hull is what is needed. Roughly, these algorithms construct the Convex Hull of a set of 2-dimensional points in $O(n \log n)$ time, which ought to be more than practically feasible for most practical top scorer lists. Obviously, my proposed algorithm computes the Upper Hull several times, but still, as the longitude of such lists are manageable, any modern computer should be able to run the algorithm in seconds.

4.2. Characteristics of the re-ranking algorithm

The proposed algorithm has some obvious properties. As a consequence, formal proofs are omitted:

- (1) The top scorer will always end up the top scorer.
- (2) The effect of the re-ranking algorithm is minimal if there is a close causality/correlation between m and g .
- (3) The effect of the re-ranking algorithm is maximal if there is no (or limited) causality/correlation between m and g .

Point 1. above states that the original top scorer, the player with largest aggregate goal score, will remain on the top of the list if Algorithm 1 is used. In the first pass of the algorithm, the first identified S' must contain this player. He has scored most goals, and hence he must have the largest g (Goals)-coordinate. Then, being first originally, will keep him first as there are no individual resorts within the Upper Hulls.

The fact that the proposed algorithm keeps original winners as that, may be seen as some kind of logical robustness.

Figure 6 explains the meaning of points 2 and 3 above. To the left in Fig-

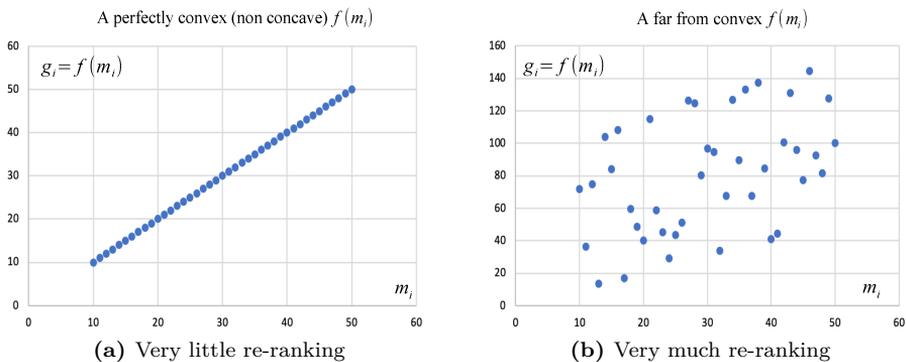


Figure 6. Explaining re-ranking behaviour for different structural data sets.

ure 6(a), a perfectly linear causality/correlation exists between m_i and g_i . Obviously, in such a situation, the first Upper Hull will contain all points in the plane

or players on the list, and the algorithm will not re-rank any players. To the right in Figure 6(b), a far less clear causality/correlation is identified, and a limited set of points will be included in the first Upper Hull, leading to significant re-ranking through the execution of the algorithm. Mathematically, and an efficient way of putting it could be:

Given a linear $f(m_i)$. Then it is always possible to describe any observed set of goals scorers and matches as:

$$g_i = f(m_i) + \epsilon_i, \quad \epsilon_i \sim ID^3(0, \sigma_\epsilon^2).$$

Then (thinking in a standard linear regression fashion)

- (1) if $\sigma_\epsilon^2 = 0$, No re-ranking takes place;
- (2) else, with increasing σ_ϵ^2 , re-ranking increases.

4.3. Applying the algorithm to academic scholar ranking

The method and algorithm presented in previous sections is surely not restricted to top scorer lists in football. Actually, the concept may be seen as a (sensible) method to sort or rank an objective containing two non weighted objectives, given that two assumptions on preference are acceptable. As such, it can have numerous alternative applications. As discussed (briefly) in Section 1, sorting of academic scholars may be an interesting alternative.

Most sensible researchers would probably object to the whole idea of trying to rank academic professionals based on a single number - like the h -index. After all, research might sometimes be a quite complex matter. A matter which should be hard to judge quality-wise, both 'today' and perhaps especially 'tomorrow'. In fact, most serious scholars would probably agree to a statement like: "*The only way to judge individual research quality is to read and (hopefully) understand the collected research articles belonging to the researcher at hand.*"

Unfortunately, our time of NPM (New Public Management) has lead to an increasing weight on simple quantitative measures, where the h -index may be seen as a "good" example, [12] may enlighten. Let me introduce a quick example⁴. Suppose two scholars A and B (at a certain point in time) are characterized by the following research production. Both have published 10 journal refereed articles. A has made some real scientific breakthroughs through two articles both receiving numerous citations, say M_1 and M_2 . However, the remaining 8 articles turned out unpopular or even unreadable and received only 1 citation each. Scholar B has also published 10 articles. He (or she) has however obtained a much more even citation profile of 10 citations for each article. The h -index of A is then 2, while B receives an h -index of 10. Hence, if the h -index is used to rank this two (obviously exaggerated) scholars, B receives a clear victory.

The point is simple. The risk of misrankings using such indexes is huge. In this case, A is loosing against B, not a good idea in my (and probably also many others) opinion. So, how could the proposed algorithm improve? Let us return to

³Identically Distributed

⁴Obviously, the algorithm will have no effect in re-ranking only two scholars. As such, this example is just for illustrative means. However, in practice, more than two players or scholars are up for ranking.

M. Berre in Table 1. Originally, he is ranked 14, but as the final re-ranking shows (see Table 6), he has moved significantly up the list, ending 4th. The reason is of course his number of matches of 452, the highest of all players, which lifts him up by the algorithmic re-rank. If we consider matches as citations, and goals as articles (just as an example), a similar effect would benefit researcher A in the above example⁵. As such, the proposed algorithmic procedure would to a much greater extent value extreme performances in one of the dimensions as opposed to the h -index which values more even performance in both dimensions. Obviously, the two given assumptions must be rephrased in this type of algorithm application; for instance something like: If both scholars have the same number of citations, the one with the largest number of articles published is the “best”. In addition, an acceptance of assumption 2; the possibility of using a combination of two scholar in outranking a third must be present.

4.4. Practical feasibility

Whether it is top scorers or top scholars, the main key to practitioners choosing to apply this algorithm is the quality of Assumptions 1 and 2. Probably, the somewhat theoretic construct of a convex combination of players (or scholars) may be criticized. Most researchers, at least those with a background in mathematics would perhaps not find the assumption especially controversial. However, it seems reasonable to assume that some might question it. From a practical implementation point of view, the top scorer-application seems easier to implemented in practice, as it is merely a matter of resorting already existing lists.

However, using the proposed method in scholar rankings may be slightly more difficult. One nice thing about the h -index, is that it can be calculated in one environment (like physics) and ‘transported’ into another (say mathematics) as it holds a numerical value⁶. The proposed algorithm does not produce any values, just lists. And then in practice, the question of what list becomes important. Are you ranked first among all economists or all sports economists? This implies some added complexity, but in the end it is the quality of the ranking that should matter. At least, as discussed above, the suggested method has nice properties when it comes to obvious shortcomings of the h -index.

Finally, be aware that the proposed algorithm takes an existing (typically sorted traditionally by goal score) list as the intended input. If one wants to apply this algorithm from scratch, one could of course stick to most goals within each Upper Hull as “internal Upper Hull” ranking. However, if one starts without any reasonable list, outliers with either very few goals (O_1 in Figure 7) in very few matches or very few goals in very many matches (O_2 in Figure 7) will typically be a part of the initial Upper Hull, and hence be high on the final list. In such cases, some kind of balancing constraint must be added. A simple possibility solving

⁵Of course, an underlying assumption that $M_1 + M_2 + 8 \gg 100$ exists.

⁶Surely, most academics would probably make strong warnings about such a strategy, as there are quite significant differences in publication frequency as well as citation frequency between academic disciplines. However, casual NPM-empiric’s shows us that academic bureaucrats occasionally may be tempted to compare apples and pears.

most of such problems could be to rule out all list candidates scoring less than the average goal score or playing less than the average match count.

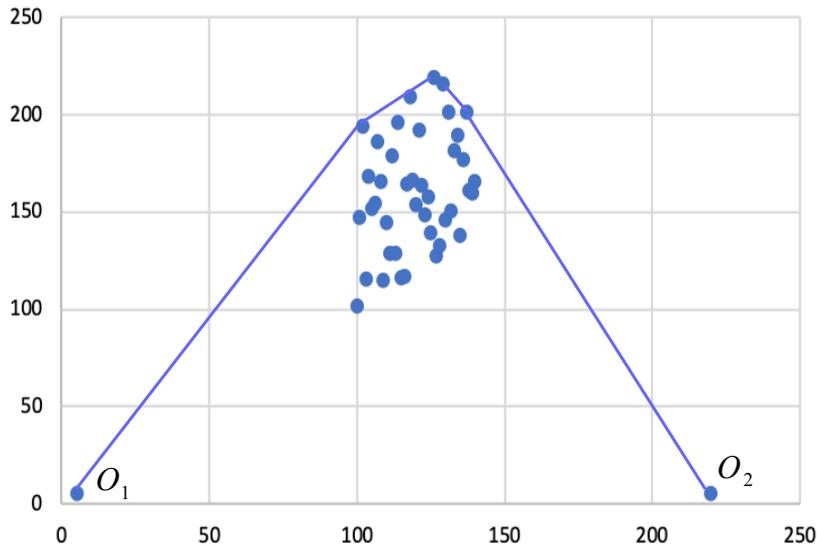


Figure 7. An example of potential problems if the algorithm is used from scratch.

However, as the title of the paper should indicate, the proposed algorithm is meant for resorting purposes, not for sorting from scratch.

REFERENCES

- [1] A. M. Andrew, *Another efficient algorithm for convex hulls in two dimensions*, Information Processing Letters **9** (1979), 216–219.
- [2] D. Avis, D. Bremner and R. Seidel, *How good are convex hull algorithms?*, Computational Geometry **7** (1997), 265–301.
- [3] C. B. Barber, A. P. Dobkin and H. Huhdanpaa, *The quickhull algorithm for convex hulls*, ACM Transactions on Mathematical Software **22** (1996), 469–483.
- [4] M. De Berg, O. Cheong, M. van Kreveld and M. Overmars, *Computational Geometry: Algorithms and Applications*, Springer-Verlag, 2008.
- [5] S. Ekern, *Time dominance efficiency analysis*, Journal of Finance **36** (1981), 1023–1033.
- [6] J. Hadar and W. R. Russell, *Rules for ordering uncertain prospects*, American Economic Review **59** (1969), 25–34.
- [7] K. K. Haugen, *Always Change a Winning Team*, Tapir Academic Publishers, Trondheim, Norway, 2012.
- [8] J. E. Hirsch, *An index to quantify an individual’s scientific research output*, Proceedings of the National Academy of Sciences **102** (2005), 16569–16572.
- [9] R. Hoffmann, L. C. Ging and B. Ramasamy, *The socio-economic determinants of international soccer performance*, Journal of Applied Economics **5** (2002), 253–272.
- [10] D. E. Knuth, *Axioms and Hulls*, Lecture Notes in Computer Science **606**, Springer-Verlag, Heidelberg, 1992.
- [11] H. Markowitz, *The optimization of a quadratic function subject to linear constraints*, Naval Research Logistics Quarterly **3** (1956), 111–133.
- [12] N. Oswald, *Does your h-index measure up?*, <https://bitesizebio.com/13614/does-your-h-index-measure-up/>, February, 2018.

- [13] E. Triantaphyllou, *Multi-criteria Decision Making Methods: A Comparative Study*, Springer-Verlag, Berlin, 2000.
- [14] J. Wesson, *The Science of Soccer*, IoP: Institute of Physics Publishing, Bristol and Philadelphia, 2002.
- [15] *List of European Cup and UEFA Champions League top scorers*, Wikipedia, https://en.wikipedia.org/wiki/List_of_European_Cup_and_UEFA_Champions_League_top_scorers, September 2018.
- [16] *Category: Sports world rankings*, Wikipedia, https://en.wikipedia.org/wiki/Category:Sports_world_rankings, October 2018.
- [17] *10 players with the best goals to games ratio ever*, Sportskeeda, <https://www.sportskeeda.com/football/10-players-best-goals-games-ratio-ever>, October 2018.
- [18] *Official! The 20 players with the best minutes-per-goal ratio in Premier League history*, FourFourTwo, <https://www.fourfourtwo.com/features/official-20-players-best-minutes-goal-ratio-premier-league-history>, October 2018.

Kjetil K. Haugen, Faculty of Business Administration and Social Sciences, Molde University College, Specialized University in Logistics, Molde, Norway
e-mail: kjetil.haugen@himolde.no