

University of Groningen

Endurance in speed skating

Kuper, GH; Sterken, E

Published in:
European Journal of Operational Research

DOI:
[10.1016/S0377-2217\(02\)00685-9](https://doi.org/10.1016/S0377-2217(02)00685-9)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2003

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):
Kuper, GH., & Sterken, E. (2003). Endurance in speed skating: The development of world records. *European Journal of Operational Research*, 148(2), 293-301. [https://doi.org/10.1016/S0377-2217\(02\)00685-9](https://doi.org/10.1016/S0377-2217(02)00685-9)

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.



Endurance in speed skating: The development of world records

Gerard H. Kuper^{*}, Elmer Sterken

Department of Economics, University of Groningen, P.O. Box 800, 9700 AV Groningen, The Netherlands

Abstract

We analyse the development of world records speed skating from 1893 to 2000 for both men and women. The historical data show that it is likely that the relation between skating speed and distance of the various events is non-linear and converges to a limit value. We pay special attention to technical innovations in speed skating, especially, the introduction of the klapskate in the 1996/1997 season, and its impact on the long-run limit value. We focus on endurance and we estimate lower bounds for world records given the current technological state of the art. We illustrate the forecasting properties of our model using the Salt Lake City 2002 Winter Games results.

© 2002 Elsevier Science B.V. All rights reserved.

Keywords: Forecasting; Athletic performance

1. Introduction

Many sports events are driven by the old Roman slogan *Citius, Altius, Fortis* (faster, higher, stronger). Human beings continuously explore the limits of human performance in, for instance, running, swimming, track and field, and weightlifting. Due to the revival of the Olympic Games in 1896, data on the history of world records are available for many sports. The analysis of world records through time is cumbersome though. The main problem of a world record time series is its distribution. Since there can only be improvements, world record times series move in one direction only, but are not allowed to go to zero. This complicates matters in estimation and fore-

casting as, for instance, is illustrated by Smith (1988). Smith (1988) tries to solve this problem by using various distribution functions, but does not find robust results. Moreover, in many cases there seems to be a linear trend in the data, which might be appropriate in the short run, but cannot be true in the long run. Most time series of world records seem to show linear trends for subintervals, which troubles identification of non-linearity as is shown by Ballerini and Resnick (1987). Thirdly, some events are characterised by technical innovations, which shift the human production function outward.

In this paper we analyse a data set that seems to suffer from all three problems (non-normality of residuals, presence of linear trends, and jumps). The data refer to long track speed skating. Speed skating records are recorded since 1893 for men and since 1930 for women. For men, world record speed skating times are available for 500 m (the shortest event) up to 10 K events. For women, we

^{*} Corresponding author.

E-mail addresses: g.h.kuper@eco.rug.nl (G.H. Kuper), e.sterken@eco.rug.nl (E. Sterken).

have world record speed skating times for the 500 m event up to 5 K (the longest official distance in women speed skating).

There have been a number of major technological innovations in speed skating. The next section briefly describes the history in speed skating with a special focus on these innovations. The major innovations are the introduction of refrigerated ovals in the late 1950s, the methods of ice preparation since the 1960s, the introduction of tight-fit clothing in the 1970s, the construction of indoor ovals since the mid-1980s, and the introduction of the klapskate since the late 1990s. In this paper we address the impact of these innovations on world records. The goal of the paper is to present estimates of lower bounds of record speed skating times conditional on technological development. These estimates are presented in Section 4. Before presenting these results, Section 3 discusses modelling issues. Section 5 concludes.

2. Innovations in speed skating

Before we proceed to analyse the data, we briefly describe the development of speed skating. The establishment of the International Skating Union (ISU) in 1892 led to a systematic registration of world records on metric distances rather than English miles. Obviously, the creation of this institution has been important to promote speed skating and initiate data collection. However, technical innovations, improving training methods, and more frequent occurrence (and perhaps earlier recognition) of exceptional athletic skills really boosted world record speed skating times. Improving training methods and fostering athletic skills result in an increased production of power, whereas technical innovations aim at reducing loss of power by reducing ice and air resistance.

As far as technical innovations are concerned we can make a distinction between innovations of the ovals, the methods of ice preparation, improved clothing, and developments to the skates themselves. Already at the end of the 19th century there was a technical innovation of the skates. Norwegian Axel Paulsen introduced lighter metal tubes and longer and thinner blades without sac-

rificing the strength of the skate. Another type of innovation is the construction of skating ovals at higher altitudes, starting with the Davos track that was founded at the end of the 19th century. Probably the most famous example in this class is the high altitude Medeo rink, where many world records have been produced until its demise in the early 1990s. The third technological innovation is the construction of refrigerated ovals. The first one was opened in 1958 in Gothenburg. The fast Medeo rink was refrigerated since 1972. Also the preparation of ice became a handicraft. In the 1960s, the ice in the Bislett stadium in Oslo was prepared with a spray of tiny droplets of water frozen in place, which resulted in a smaller area of contact with the skate blade and thus less friction, at least until the droplets were worn down. Skating clothes were improved also. The Swiss skating veteran Franz Krienbühl introduced the tight-fit suits in 1976. In recent years some skaters experiment with special sharkskin suits which are supposed to further reduce the air resistance. More important to the improvement of the world record is the construction of indoor 400 m ice rinks. The first indoor ovals were developed in Heerenveen, The Netherlands, in 1986 and in 1987 in Calgary, Canada, for the 1988 Winter Olympic Games. The Dutchman Van Ingen Schenau establishes a major innovation of the skate itself in 1996. Disconnecting the blade from the heel of the skate and placing a pivot point under the ball of the foot allows skaters to use the full extension of the leg to achieve maximum power and glide. The advantages of this skate, called the klapskate, were already known since the end of the 19th century. As we will illustrate below, this innovation led to a serious improvement of world records on all distances.

In order to model the historic development of speed skating world records, we use data published by the ISU. The ISU website¹ gives the world records since 1893 for each event. We take the world record of a year by the record of July 1st. We construct series that are normalised to the 500 m event (this is a common way of notation for the

¹ URL: <http://www.isu.org/historical/sswrecs.html>.

ISU in large championship events). So we measure seconds per 500 m for all events (both male and female records). Fig. 1 illustrates the data. Fig. 1 clearly reveals that there is serious progress in speed skating. One can also observe that certain events lack competition. For men this is the 3 K race, because the 3 K race is not organised in big skating events (like world championships, Olympic Games, or World Cup). For women the 5 K event shows no improvement from 1945 up to 1983. The reason is that only in 1983 the 5 K event returned on the official women skate calendar.

Fig. 1 shows a remarkable improvement of the world record speed skating times. Note also that

the 1000 m event yields the highest average speed. Due to the loss at the start the average speed on the 500 m is a bit lower. Here we can see the analogue with running. The 200 m dash record also yields a higher average speed than the 100 m dash record. Fig. 1 also shows convergence of speed across events.

Fig. 1 also illustrates some of the problems in modelling world records. These problems violate the normality assumption of the residuals in linear regression, which are required for statistical inference. First, the time series of world records move in one direction only. Secondly, the time series seem to show linear trends for subintervals.

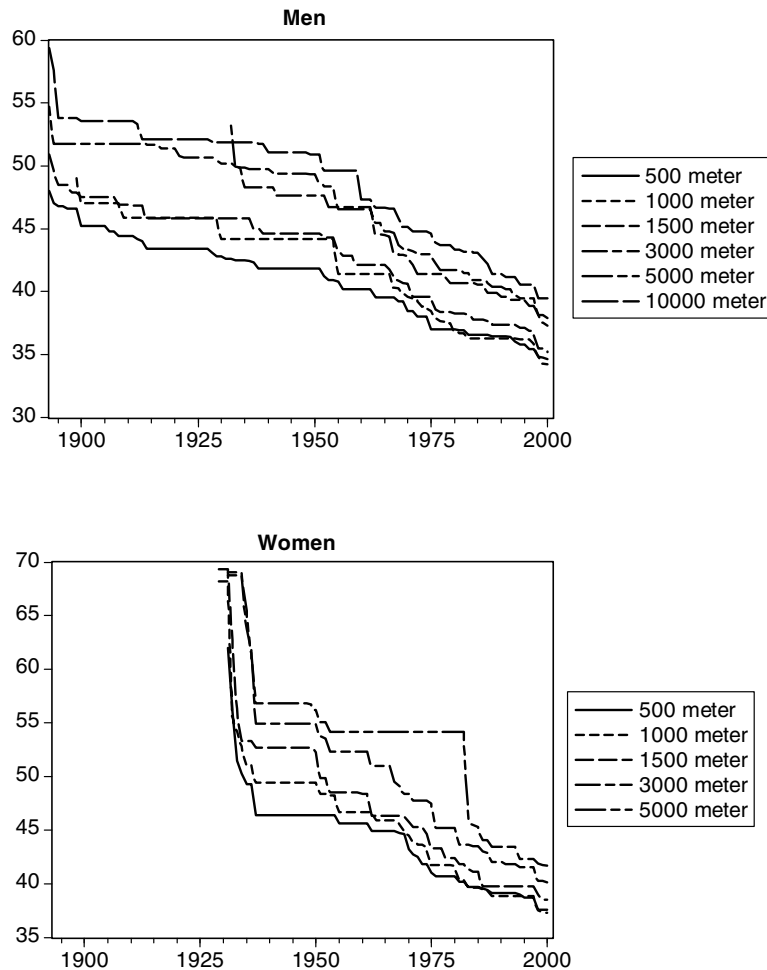


Fig. 1. Speed skate world records for men (top) and women (bottom) in seconds per 500 m.

And finally, the causes of the improvement of world records vary. Clearly we observe a jump in the data at the end of the 20th century which is caused by one specific technical innovation, in this case the introduction of the klapskate.

In the remainder of the paper we focus on endurance. It is well known that world records of sprint events satisfy different statistical properties, due to the physiological characteristics of the human body. To give some idea of the skating speed at long distances, we show some information about skating speed in marathon events in Fig. 2. Except for difference in distances, these events are different from speed skating. Marathons are mass start races and are either indoor or outdoor on non-refrigerated ice. Men typically race 100 and 150 laps on ovals (about 385 m per lap) and 100 and 200 K on natural ice. Women race 50 and 60 laps indoor and 60 and 200 K outdoors. The most well known and longest event is the 200 K Elfstedentocht. Because of lack of ice in The Netherlands this race is organised every year in Austria and Finland. There are no official world records for marathons so these data merely serve as an illustration and will not be used in the analyses in the next sections.

Fig. 2 shows that it is likely that the relation between skating speed and distance of the event is non-linear and converges to a limit value. It is the development of the long-run limit value of skating speed that draws our attention in the remainder of the paper.

3. Modelling world records

In this paper we analyse endurance of the human body in skating long-run distances. We express performance by the average speed to skate certain distances. So it is required to come up with theory on the determinants of skating speed. Basically, one can distinguish between two main classes of speed skating events on oval rinks. First, one can analyse sprint events, which is anaerobic skating. Here the muscular structure of the body is one of the key determinants of speed. Also the reaction time at the start is important. The physiologic differences between women and men determine to a large extent the differences in results between the sexes. It is also obvious that in sprint events there are large differences in speed during the event. This fact makes it troublesome to use average speed during an event as an indicator of performance.

Secondly, one can analyse the aerobic events, where endurance of the body is the most prominent physical characteristic. Here it is the maximum oxygen intake (the so-called $VO_2\text{max}$) that plays a key role. Again the muscular structure differences between women and men make that men are able to skate faster in absolute terms. Once we analyse velocities across distances the differences between men and women are less prominent (see hereafter). The aerobic events are characterised by the fact that a constant velocity is optimal (Keller, 1974).

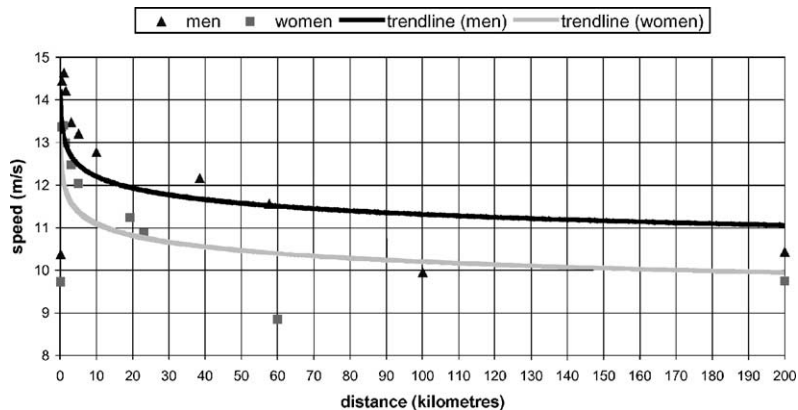


Fig. 2. Speed in long track events (≤ 10 K) and marathon events (>10 K) for men and women (smooth lines are power trends).

In this paper we analyse lower bounds of aerobic events, like Blest (1996) has done for track and field. Especially, we focus on the lower bounds of speed skating time. By doing so, we are interested in the human production frontier. This implies that we analyse skating data given perfect race conditions, the use of the best training methods and equipment possible, and perhaps more important, the mental ability of the athlete to compete. We cannot analyse behaviour worse than the frontier, apart from saying that the outcomes for all competitors are probably normally distributed for a certain age. We are also interested in the impact of distance on velocity. It is clear that anaerobic events involve higher speeds. We concentrate on statistical evidence on the results. Riegel (1981) and Blest (1996) use a log-linear model to relate time used to cover a distance with the distance itself:

$$\log(t) = a + b \log(\text{dist}) \quad (1)$$

to fit world records, where t is the time in seconds and dist , the distance in meters. If b exceeds unity, time increases more than linearly with distance. This model can also be rewritten in the terms of speed, using $tv = \text{dist}$:

$$\log(v) = -a + (1 - b) \log(\text{dist}), \quad (2)$$

where v is the velocity in meters per second. Least squares estimates for model (1) or (2) using data on speed skating world records for the period 1932–2000 are $b = 0.6$ for men, and $b = 0.8$ for women. It is important to note that the time used in an event increases less than linearly with distance.

In the remainder of the paper, we will not use models (1) or (2), since they lack a proper definition of the long run. We focus on endurance and use a model based on the one proposed by Francis (1943):

$$v = C + A/(\log(\text{dist}) - B),$$

where C is the speed at very long distances (i.e., if dist approaches $\exp(B)$ in the limit), $\exp(B)$ is the asymptote of the distance where the maximum speed can be observed and A is a measure of the decrease of speed as distance increases. Mosteller and Tuckey (1977) use the following shifted power transform variant:

$$v = A(\text{dist} - B)^x + C \quad (3)$$

to linearise the Francis-model, where x is a negative real number. B is the distance at which the race changes from an anaerobic to an aerobic event (say from sprint to long-distance race). In the literature on track and field there are estimates of B . Keller (1974) estimates B to be 291 m. If we simply look at the total time used, a reasonable estimate for the switch from anaerobic to aerobic events for long distance speed skating, would be around 700 m.

The Francis model is an interesting tool to fit the progress of world records, which is the focus of our contribution. Suppose we can estimate C for various moments of time and construct a time series $C(t)$. Again, $C(t)$ represents the skating speed at very long distances. In the literature there are a few parametric forms known to model the development of $C(t)$. For our goal we want a non-linear relation with a clearly defined positive limit value. To that extent we can apply the Chapman–Richards form:

$$C(t) = d - a[1 - \exp\{-b(t - \text{baseyear})\}]^g. \quad (4)$$

In this paper, we focus our attention on the development of the long-run limit speed, i.e., the limit of $C(t)$. Because the limit of C equals $d - a$, the parameters of interest in Eq. (4) are d and a . Furthermore, parameter b indicates the speed at which $C(t)$ approaches its limit, $(\log 2)/b$ is the half-life, i.e., the time it takes to advance to half the ultimate speed.

4. Results

In this section we model the development of the world records on various distances in speed skating. As argued above we apply the Francis model and the Chapman–Richards approach to model the dynamics of the long run average speed. The analytical problem is that the causes of the improvement of world records vary. It might be technical innovation, better training, exceptional personal skills, and doping. It is typically hard to disentangle all possible causes. Despite that we focus on technological innovations.

We can model the theoretical speed in skating as follows (see Fig. 3). We model skating speed as a function of inputs, like training and medical care. It is likely that there are diminishing returns to inputs. Depending on the distribution of talent it is likely that once in a while a talented young skater will reach a higher point on the curve. A famous example is Eric Heiden who is a three-time overall world champion (1977–1979). Furthermore, he won all five men's gold medals at 1980 Olympics, and set records in each event. The production frontier is a function of time t . As time passes the probability of hitting a higher point on the same locus will increase. This represents a kind of automatic progress. There is big change of hitting the frontier if there is a major innovation. In that case the whole curve shifts upward. Note that a vertical jump to a higher production frontier will also increase marginal productivity, because the tangent slope of the frontier, representing productivity, is steeper.

We estimate speed in skating in two stages. First, we estimate a pooled version of Eq. (3). The cross sections are the years and the sample size is determined by the number of different speed skating events. In estimating the pooled model we put the shift from anaerobic to aerobic events (the unidentified parameter B) equal to 700. This choice

is based on other results in track and field (see Section 3). The model is completely identified if we also choose a value for x . We use after experimenting $x = -0.267$ (see also Grubb, 1998). We use a uniform A (constant over time), but estimate C to be time dependent (fixed year-effect). By picking a constant A we assume a constant loss rate, which has also been found in other endurance sports (see Grubb, 1998). The sample starts in 1932 for men and women (we exclude the earlier data for men, due to the low variability). Parameter A for men is 13.053 and for women 11.447. As is indicated above, C represents the speed that can be used at long distances. Fig. 4 gives the time series of the C 's for men and women.

One can observe from Fig. 4 that the long-run speed for men is developing at a fixed rate from 1955 up to 1997. For women there are more episodes of shocks (we excluded the 5 K event for its discrete character). Fig. 4 also clearly shows that the klapskate is a real innovation for both men and women. One can also observe that the other innovations, as mentioned in Section 2, are visible in the data. However, these innovations are less prominent than the shock due to the klapskate.

Next we estimate Eq. (4) for both men and women. Table 1 gives the estimation results. We use the sample 1960–1996. Note that we start in

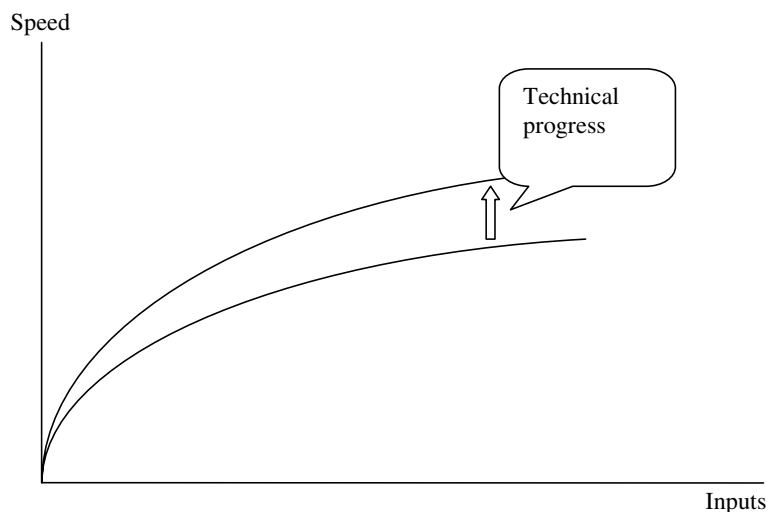


Fig. 3. Speed and technical progress.

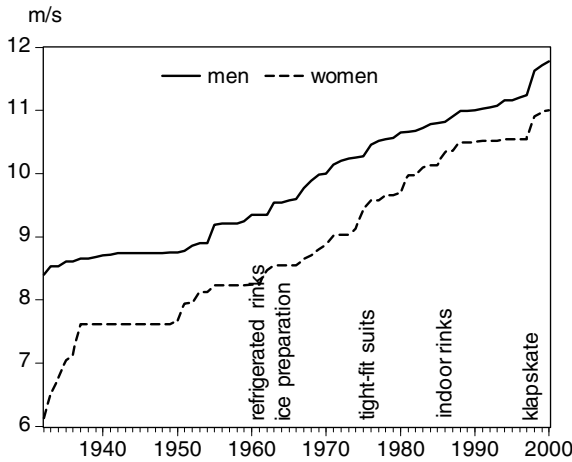


Fig. 4. Long-run speed for men and women.

Table 1
Long run speed using FIML (*t*-values between brackets)

Model: $C(t) = d - a[1 - \exp\{-b(t - \text{baseyear})\}]^{g^a}$		
	Men	Women
<i>d</i>	9.307 (285.65)	8.462 (343.07)
<i>a</i>	-2.230 (-17.433)	-2.395 (-20.486)
<i>b</i>	0.059 (7.787)	0.103 (7.748)
<i>g</i>	1.429 (9.831)	3.621 (5.546)
SSR	0.044	0.081
Implied <i>d - a</i>	11.537	10.857
Half-life	11.7 years	6.7 years
Sample	1960–1996	1960–1996

^a $C(t)$ represents the speed at very long distances, and limit of C is given by $d - a$.

1960 due to the low variability of the series during the war and the high volatility and catching up short after World War II. The sample ends in 1996 in order to be able to analyse the effects of the klapskate. From these results we can calculate the limit speed $d - a$. These calculations show that men can reach a limit speed of 11.54 m/second, while women are able to reach 10.86 m/second in the very long run. The half-life for men is $\log(2)/0.059 = 11.7$ years and for women $\log(2)/0.103 = 6.7$ years.

Estimating the C -function could also provide some insights into the remarkable performances of famous speedskaters. Is it the case that Ard Schenk, Kees Verkerk, Fred Anton Maier, Eric Heiden, or Johan Olav Koss were extreme outliers? If they were, how can they be compared? We analysed the residuals from the C -function and tested for each year their significance. The problem in doing so is that we cannot identify whether the quality of the skater or a technological innovation is driving the outlier. Our results (not reported here, but available upon request) show that we can only accept an outlier for the period 1967–1972, the years that Ard Schenk, Kees Verkerk and Fred Anton Maier were ruling the skating world. From this perspective our model shows that their performances are better than the relative performances of Heiden or Koss. It is good to note that our model is not well suited to compare athletic performances (like the extreme value approach by Robinson and Tawn, 1995, or the athletic performance comparison by Grubb, 1998).

For men and women we observe a jump in C from 1997 to 1998 of about 0.3. This implies that the gain in speed of the klapskate is about 30 cm per second in only one year. In the long run we expect the gain of the klapskate to be about 60 cm per second. Again, this assumption is based on the theoretical notion of Fig. 1, where a vertical jump to a higher frontier also implies a higher rate of progress. The assumption of 60 cm per second is equivalent to a gain of about 1 second per lap at all distances (for the 10 K this implies a gain of 25 seconds), which is what generally is believed to be the gain of the klapskate. These estimates are based on Van Ingen Schenau et al. (1997). So we identify this innovation as a shift from the frontier, as explained in Fig. 1.

Finally, using the model we are able to forecast the ultimate world records. For all distances, except the shortest distance, we calculate the world records by

$$v_{\infty} = A(\text{dist} - 700)^x + (d - a) + 0.6.$$

The results are given in Table 2. Compared with current (April 19, 2001) world record times we can draw a number of conclusions. First, the expected gain is expected to be the largest in women events.

Table 2
Predicted long-run world records

	Current 19-04-2001	Predictions		
		With klapskate	Expected gain (%)	Without klapskate
<i>Men</i>				
500	34,32	NA	–	NA
1000	1.08,28	1.06	2.4	1.09
1500	1.45,20	1.44	0.5	1.49
3000	3.42,75	3.37	2.4	3.47
5000	6.18,72	6.09	2.5	6.26
10000	13.03,40	12.33	4.0	13.08
<i>Women</i>				
500	37,29	NA	–	NA
1000	1.14,13	1.11	3.4	1.14
1500	1.54,38	1.52	2.0	1.57
3000	3.59,27	3.52	2.9	4.03
5000	6.52,44	6.34	4.6	6.53
10000	NA	13.22	–	14.03

Second, the expected gain is more pronounced for the longest distances for both men and women. Furthermore, we expect that the men's current 1500 m world record has a low probability to be improved significantly in the near future. Table 2 also shows the ultimate world records if the traditional fixed-blade skates would have been used. Again, the results indicate the enormous impact of the klapskate. Without exception, the current world record times exceed the hypothetical ultimate, fixed-blades, world record times.

Using our model, we can simulate the ultimate world record for the women's 10 K. According to our estimates, a time of 13 minutes and 22 seconds is ultimately possible.

A natural question to ask is when these ultimate times will be reached. The answer must be: In the very long run. However, using the information on half-lives we can make a series of forecasts. The first column of Table 3 gives the current world record times and the last column repeats our ultimate, long-run forecasts. The intermediate forecasts are based on the estimated half-lives (11.7 and 6.7 years for men and women, respectively) and the "quart"-lives (4.9 and 2.8 years for men and women, respectively).

A final check on the results is an analysis of the performance at the Olympic Winter Games in Salt

Table 3
Predicted world records

	Current 19-04-2001	Predictions		
		2006	2013	Ultimate
<i>Men</i>				
1000	1.08,28	1.07	1.07	1.06
1500	1.45,20	1.45	1.44	1.44
3000	3.42,75	3.41	3.40	3.37
5000	6.18,72	6.16	6.13	6.09
10000	13.03,40	12.55	12.48	12.33
		2004	2008	
<i>Women</i>				
1000	1.14,13	1.13	1.12	1.11
1500	1.54,38	1.53	1.53	1.52
3000	3.59,27	3.57	3.55	3.52
5000	6.52,44	6.47	6.43	6.34

Lake City in February 2002. We predicted the winning times, knowing that the Salt Lake City speedskating track is a state of the art track that allows for world best performances. Moreover, athletes were considered to be optimally prepared for the Olympic performance. Table 4 gives the main findings.

We report the predictions (as of April 2001), the realisations and the relative absolute error (RAE) for all events except for the 500 m. Table 4 shows that in some cases the model predicted the winning new world record times rather well. The men's 1500 m and 5 K forecasts seem to give the worst

Table 4
Predictions and realisations of Salt Lake City 2002 results

	Predictions	Realisations	RAE
<i>Men</i>			
1000	1.07	1.07,18	0.0027
1500	1.45	1.43,95	0.0101
5000	6.17	6.14,66	0.0062
10000	12.57	12.58,92	0.0025
		MRAE	0.0054
<i>Women</i>			
1000	1.14	1.13,84	0.0023
1500	1.54	1.54,02	0.0002
3000	3.58	3.57,70	0.0013
5000	6.49	6.46,91	0.0051
		MRAE	0.0022

results. For women we find a better approximation than for men (0.22% (mean) relative absolute error (MRAE) for women versus 0.54% for men).

5. Conclusion

Speed skating is characterised by important technical innovations. Especially the introduction of the klapskate in the 1996/1997 season dramatically improved world records speed skating, both for men and women. This development led us to analyse the development of world record speed skating times since the end of the 19th century. Assuming no new major innovations, which are difficult if not impossible to predict, and focusing on endurance we analyse and predict world records using models that have been used in track and field.

Focusing on endurance implies that we decompose velocity into two terms. The first term is velocity in very long distance (endurance) events and the second term is event-specific and constant over time. Endurance speed increases over time and is modelled for the pre-klapskate period. The klapskate effect is added as a constant. Our predictions point at a future improvement of the 10 K men's world record of 30 seconds and an improve-

ment of the 5 K women's world record of about 20 seconds. We illustrate the working of the model in analysing the predictions for the Salt Lake City Winter Games 2002.

References

- Ballerini, R., Resnick, S.I., 1987. Records in the Presence of a Linear Trend. *Advances in Applied Probability* 19, 801–828.
- Blest, D.C., 1996. Lower Bounds for Athletic Performances. *The Statistician* 45, 243–253.
- Francis, A.W., 1943. Running Records. *Science* 98, 315–316.
- Grubb, H.J., 1998. Models for Comparing Athletic Performances. *The Statistician* 47, 509–521.
- Keller, J.B., 1974. Optimal Velocity in a Race. *American Mathematics Monthly* 81, 474–480.
- Mosteller, F., Tuckey, J.W., 1977. *Data Analyses and Regression*. Addison-Wesley, Reading, MA.
- Riegel, P.S., 1981. Athletic Records and Human Endurance. *American Scientist* 69, 285–290.
- Robinson, M.E., Tawn, J.A., 1995. Statistics for Exceptional Athletic Performances. *Applied Statistics* 44, 499–511.
- Smith, R.L., 1988. Forecasting Records by Maximum Likelihood. *Journal of the American Statistical Association* 83 (402), 331–338.
- Van Ingen Schenau, G.J., Houdijk, H., Koning, J.de., 1997. Backgrounds of the Klapskate. <http://www.gironet.nl/home/cvs/taave/schaatsen/article/klapskate.htm>.