An improvement index to quantify the evolution of performance in field events

HAAKE, Steve <http://orcid.org/0000-0002-4449-6680>, JAMES, David <http://orcid.org/0000-0002-1135-626X> and FOSTER, Leon <http://orcid.org/0000-0002-1551-0316>

Published version

HAAKE, Steve, JAMES, David and FOSTER, Leon (2014). An improvement index to quantify the evolution of performance in field events. Journal of Sports Sciences, 33 (3), 255-267.

Copyright and re-use policy

An improvement index to quantify the evolution of performance in field events

Journal:	*Journal of Sports Sciences*
Manuscript ID:	RJSP-2013-1125.R1
Manuscript Type:	Original Manuscript
Keywords:	performance index, non-linear regression, drugs testing

An improvement index to quantify the evolution of performance in field events

1. Introduction

Athletic performance has increased rapidly over the last 120 years since the creation of the Modern Olympic Games with an apparent evolution of performance much faster than that of natural human evolution. The reasons for this growth have been attributed to common themes such as globalisation, population growth, technology, drugs and coaching interventions (Denny 2008; Ernst and Simon 2013; Foster *et al.*, 2011; Lippi *et al.*, 2008). Balmer *et al.* (2012) went further and searched for specific interventions, managing to quantify the introduction of the Fosbury Flop and, in part, the introduction of new pole-vaults.

An improvement in performance can come from two sources: (1) an improvement in the efficiency of the athlete; or (2) through a reduction in energy losses. In the high jump, for example, the former is associated with the introduction of the Fosbury Flop technique while, in the pole vault, the latter is associated with improvements in the pole design (Burgess, 1996).

Previous research has tended to focus on natural limits to performance progression allied to discussion on the causes of performance improvement (Lippi, 2008; Foster *et al.*, 2010; Balmer *et al.*, 2012, Haake *et al.* 2013). Research by Haake and colleagues (Foster 2012; Foster *et al.* 2010; Haake 2009; Haake *et al.* 2013) tried to identify how much of the improvements seen in sport are due specifically to technology. Haake (2009) showed that yearly results in field events could be used to analyse performances by equating the recorded heights and distances to useful work done. In the high jump or pole-vault, the useful work done was equated to the potential energy of the athlete successfully clearing the bar. Using the height of the bar h as the minimum height of the jump (since the centre of mass may go over or under the bar depending upon body shape), the useful work done on the athlete's centre of mass m was given as,

$$W = mgh \tag{1}$$

where g is the acceleration due to gravity. An improvement in performance for a jump h compared to a baseline jump h_o in year d_o was found by taking their ratio to give a performance improvement index,

$$Index = 100 \times \frac{W}{W_0} = 100 \times \frac{(mgh)}{(mgh)_0} = 100 \times \frac{h}{h_0} \quad (2)$$

for a constant mass *m* with the multiplier of 100 used to give a percentage. A value greater than 100% indicated an improvement in performance for jump *h* compared to jump h_0 with a corresponding increase in useful work done. It is difficult to get the mass of athletes retrospectively but Norton & Olds (2001) showed that the height and weight of athletes is increasing. This would cause the index in (2) to be an underestimate of the increase in work done.

In throwing events, the useful work done to propel an object distance *s* through a parabola in a vacuum is proportional to the maximum potential energy in (1). This allowed Haake (2009) to replace *h* and h_0 with the throw distances *s* and s_0 in (2) for events such as the shot, discus and javelin.

The collection of performance data in field events started in the 1890s and recent research has shown that a yearly mean of the top athletic performances allows trends to be seen. The median tends not to be used as it is dominated by the resolution of the measurement system (e.g. 1 cm in the pole vault and high jump or 0.1 s in running events prior to 1976). Using the mean also reduces the effect an individual performance might have on the trend. The choice of data is dependent upon what is available at the time of collection; Berthelot *et al*. (2010) used the top 10-performances, Ernst and Simon (2013) the top-20 performances and Haake *et al*. (2013) the top 25-performances.

In the study of performance *per se*, many researchers have used exponential functions with a natural limit to reflect the natural shape of the data (Blest, 1996; Denny, 2008; Nevill and Whyte, 2005;). Balmer *et al*. (2012) used a double sigmoid fit in their search for the effect of the Fosbury flop and new pole-vaults. Haake *et al*. (2013) used a three-parameter model proposed by Ratkowsky (1983) to model the underlying growth in performance in running, adding step functions to represent 'instantaneous' changes due to rule changes and linear functions to represent gradually introduced interventions.

The aim of this paper is to use data from field events collected since the 1890s allied to mathematical modelling of the data to quantify the effect of influences and interventions on performance.

2. Methodology

Through stages are proposed: (1) the collection of performance data; (2) the conversion of the data to the performance improvement index; and (3) the fitting of models to the data to identify the causes of improvement (or decrement).

Stage 1: data collection

With institutional ethical approval, the top-25 individual performances were collected from open-source performance statistics websites (International Association of Athletics Federations (IAAF, 2011; Rabinovich, 2013) for 8 men's and 5 women's field events between 1890 and 2012. Only an athlete's top performance was used in each year and, thus, each athlete appeared only once in the yearly list.

Stage II: the performance improvement index

Haake et al. (2013) showed that any baseline date chosen in running ought to be after the 2nd World War due to the distinct reductions during 1939-1945. A baseline date of d_o = 1948 was chosen and the performance improvement index calculated using (2) and the mean of the top-25 performances in each year. All performances, therefore, have an index value of 100% in 1948.

Stage III: *Modelling of the data*

The overall secular rise in the performance improvement index at date d (in centuries from 1800) can be modelled using,

$$Index = 100 \times \left(\frac{h}{h_o}\right) = 100 \times \left(L - e^{-a_1 \cdot a_2^d}\right) \quad (3)$$

where L is the limit, and a_1 and a_2 are constants which determine the shape of the curve (Haake *et al*., 2013). The shape of this function ensures a steep initial rise with a gradual levelling off to a limit L.

For an intervention that creates a step change in performance time Δh_i, then the performance improvement index in the period after the step is given by,

$$Index_i = 100 \times \left(\frac{h + \Delta h_i}{h_0} \right) = 100 \times \left(\frac{h}{h_0} + c_i \right) \tag{4}$$

where $c_i = \Delta h_i / h$

and i=1 to n where n is the total number of step changes. Combining (3) and (4) for an exponential rise with step changes in years d_i gives,

$$Index = 100 \times \left(L - e^{-a_1 \cdot a_2^d} + c_i \right) \tag{5}$$

An intervention that takes effect over time can be modelled using a linear change given by,

$$c_i = \Delta c_i \left(d - d_{i_o} \right) \tag{6}$$

where the intervention rises by gradient Δc_i between dates d and d_{io} to a final date d_{if}.

The oscillation due to the Olympic Games found by Haake *et al.* (2013) is represented as a sine function of amplitude A so that (5) becomes,

$$Index = 100 \times \left(L - e^{-a_1 \cdot a_2^d} + c_i + A \sin\left(\omega d + \phi \right) \right) \tag{7}$$

where ω and ϕ are the frequency and phase to fix the period to 4 years such that the maximum occurs in an Olympic year.

A bespoke Matlab programme was used to carry out a non-linear least squares regression analysis using the Levenberg-Marquardt algorithm (Seber, 2003; Moré, 1978). The model in (7) was fitted to the performance improvement data for each of the 13 events: the extended exponential curve was entered first so that the minimum number of parameters allowed was 3. The step functions c_i and the sinusoid of amplitude A were then fitted in a stepwise manner, introducing the interventions in order of significance determined by the highest change in adjusted $R^{2\dagger}$ until it no longer improved.

† The adjusted R^2 takes into account the introduction of additional parameters.

Interventions to be assessed

It was assumed initially that the same interventions that were found to by Haake *et al.* (2013) to affect running would also affect field events, i.e. an underlying exponential rise in performance due to globalisation (3), a periodic influence of the Olympic Games (7), the introduction of random controlled drugs testing in 1989, and the formation of the World Anti Doping Agency (WADA) in 1999. Although the latter was formed in late 1999, it is assumed that the effect did not fully take place until 2000. Other effects commonly assumed to have affected field events were also searched for. These are as follows:

1. High Jump: Dick Fosbury introduced a new jumping technique nicknamed the 'Fosbury flop'; it has been assumed to be a key driver of performance (Balmer *et al*. 2012). A linear function was used to describe this performance improvement between the dates of 1968 and 1976 suggested by Foster (2012) as the key period for its development.
2. Pole vault: The use of composite poles made of glass fibre changed the technique of pole-vaulters to a gymnastic procedure from the mid-1950s onwards and contributed to an improvement in performance (Haake, 2009). A linear function was used to simulate the gradual increase in the use of composite poles between 1956 and 1972.
3. Javelin: The inertial characteristics of the javelin were improved with the introduction of hollow javelins between 1953 and 1956. The IAAF introduced the following rule changes: in 1986 to move the centre mass for the men's javelin; in 1992 to ban the use of roughened javelins; and in 1992 to move the centre of mass of the women's javelin. The introduction of the hollow javelin was treated as a linear uptake, while the rule changes were considered as step changes.

3. Results

Figure 2 shows the performance improvement index with the stepwise regression models and the date the interventions were introduced (*note*: the vertical scales have been optimised for each event to allow interventions to be visible). The performance improvement index reveals some general characteristics: (1) performance improves more in throwing than jumping events, and (2) performance improves more in women's than men's events. The greatest performance improvement was seen in the women's discus and shot put, while the smallest was in the men's long jump and triple jump.

The adjusted R^2 in Table 1 gives an estimate of the explained variance between the final models and the performance improvement index. It varies between 0.957 for the men's long jump and to 0.994 for the women's high jump and indicates that the models are good predictors of the data. The trends in the data for all events are largely accounted for by the exponential function, with the implication that a steep initial rise followed by an asymptotic limit is a representative shape for the overall rise in performance.

The Olympic oscillation was found to increase the adjusted R^2 in 12 events, compulsory random drugs testing in 11, the formation of WADA in 10, and the technology related interventions in 8 events (Table 1). Equations (3) to (7) and the parameters in Table 1 can be used to calculate the performance improvement index in 2012 and the change caused by each intervention with the index set to 100% in1948 (Table 2). This is explained below.

The model found for each event gives the secular rise in the data using the exponential equation in (3), combined with step and linear changes and the Olympic oscillation. A plot of the exponential function in the men's long jump in Figure 3, for example, shows how performance would have continued had there been no step change interventions in 2000. In this example, the exponential function would have risen to 110.4% had there been no other influences on performance other than a global rise. When WADA was introduced, performance dropped by 1.7%. Additionally, the Olympics caused an oscillation such that performance improved in 2012 by 0.4%. The final performance in the men's long jump was 109.1%, or a final *change* of 9.1%.

This data is tabulated in Table 2 for all events and the *changes* between 1948 and 2012 constructed in a single graphic in Figure 4. The total length of each bar in Figure 4 indicates the underlying secular change between 1948 and 2012 given by (3) (e.g. 10.4% in the example of the men's long jump). The elements to the left of the vertical axis have reduced performance while those to the right have improved performance (i.e. decreased or increased the useful work done). The remainder (8.3%

in the case of the men's long jump) is shown by the white bars in Figure 4 and represents the global effects that have improved the athletes between the two dates (e.g. improved nutrition, sports science support and population increase.)

Common characteristics

Inspection of Figure 4 and Table 2 again shows that women's field events have improved more than men's events in 2012 compared to 1948 (139.4% for women compared to 125.6% for men). The performance improvement index was also greater in throwing than jumping events increasing by 140.9% and 122.7% respectively.

The introduction of drugs testing in 1989 (Table 2 (iii)) and caused performance to drop by a mean of -4.7% (-6.6% for women and -2.8% for men). The effect was greater in throwing events than jumping events with mean drops of -7.0% and -1.6% respectively. The formation of WADA showed a smaller mean reduction overall (-2.5%) with similar trends to drugs testing when men are compared with women and jumping with throwing.

The Olympic Games oscillation showed a smaller explained variance than the drugs interventions, the effect of which was to cause a slight increase in the performance improvement index in 2012 (and corresponding Olympic years) of around 0.5% for field events (0.6% for women and 0.4% for men).

Technology

There were 5 interventions in the 'technology' category that caused an increase in performance in field events (Table 2 (vi)). The uptake of the Fosbury flop in the high jump improved performance by 1.7% for men and 4.8% for women, while the use of composite poles in the men's pole vault increased performance by 7.9%. The use of hollow javelins between 1953 and 1956 showed a similar increase for men and women of 5.8% and 4.3% respectively. The mean increase in performance by all positive effects in field events was 4.9%.

The 3 rule changes by the IAAF to affect the flight of the javelin caused reductions in performance in the javelin: moving the centre of mass in the javelin reduced performance by -10.5% for men and -9.0% for women while the rule on tail roughness reduced the women's event by -5.3% (this was not found in the men's event).

4. Discussion

The aim of this paper is to quantify the effect of influences and interventions on performance in field events. The performance improvement index quantifies the change in useful work done by a cohort of athletes over time, with the resulting change a summation of the influences and interventions between two dates. Improvements in performance and the usefully available work come about though improvements to the athlete cohort and reductions in energy losses. The modelling techniques used here have managed to identify the individual inputs to cause these changes in performance.

The results show that performance has changed most in the throwing events and least in the jumping events. The reasons for this might be that there was less depth of competition in throwing events in 1948 and strength and conditioning strategies allowed throwing to improve more. An additional temporary effect was the use of drugs since the introduction of drugs testing and WADA decreased performance significantly. Since the data analysed was the mean of the top 25 performances in each year then a majority of athletes appear to have been culpable in the use of drugs prior to drugs testing. The corollary must also be true, i.e. that drugs testing has been largely successful with the implication that a majority of athletes are not using performance enhancing drugs (or that their ability to do so has been seriously limited).

Rule changes by the governing bodies of sport are usually introduced as a remedy to an undesirable situation. Examples of this are the introduction of the two rules in javelin to limits tail roughness and to move the centre of mass forward; these reduced performance at between -5.3% and -10.5%. Flexibility in the rules, however, can also enable technology to improve performance: hollow javelins improved performance in the men's javelin by 6% while the introduction of composite poles increased performance by 8%. While the rule changes to javelins created a greater drop in performance than the hollow javelins used to improve it, the improvements from composite poles were accepted and are still in use today.

The remaining causes of improvement have been lumped together into a 'residual global effect' that groups together all other influences such as coaching, training, nutrition, sports science and population increase. This represents the majority of the rise for most field events at between 8% and 50% (for the men's long jump and women's discus respectively). Norton & Olds (2001) showed that the anthropometrics of athletes have become more extreme compared to the general population and elite athletes are becoming outliers in the population distribution. Allied to this, the global population has also increased from around 2.5 in 1948 to 7 billion in 2012, which increases the likelihood of outstanding athletes being found (Yang, 1975). Thus, a large proportion of the residual effects is talent identification of specialised athletes with specific body shapes necessary for success at the elite level. This is then followed by athlete optimisation through coaching, sports science and nutrition, although the method here cannot separate out the different influences.

The graphs in Figure 2 show that performance reaches a natural limit unless an intervention stops it from doing so. Most sports appear to be reaching a plateau and the dates when field events will reach 99.9% of their model limits *L* are shown in Table 1. All throwing events, except the javelin, have already reached their limit while jumping events are predicted to reach it by the mid 2030s. In track events, Haake *et al*. (2013) showed that only sprint events are yet to reach their limit with all events over 400 m having already reached their limit.

The close association between sprint events and the long and triple jumps is evident in the data in Table 2 with increases in performance of similar magnitudes. As with field events, women's running events improved more than men's events and those sports in which the initial athlete population is likely to have been small and relatively uncompetitive tend to see the largest improvements in performance. Events starting at a lower performance level in 1948 would automatically show greater improvement than more established events. Long distance running showed improvements caused by an influx of a new population (African athletes). This sort of effect was not found in any field event although it might have occurred in sports such as the javelin, for instance, which has long been dominated by northern European countries.

Denny (2008) showed that greyhounds and racehorses reached a limit in performance when selective breeding was unable to improve the population further. Unless there is a cause to change performance, then it will naturally reach equilibrium where the world's best athletes, techniques or technologies are already in use. Going back in time, performances prior to the 1880s (Figure 1) must have had a previous equilibrium value, which reflected the coaching methods, techniques and abilities of the athletes and the relative lack of athletic competition at that time.

Looking at performance over an even longer period of time of, say, a thousand years, the overall jump since the 1890s and current levelling off might seem like any of the transient interventions described in this paper. One observation, then, is that the global increases in population, health and prosperity brought on by the industrial revolution of the 19th Century was probably a large-scale intervention in its own right. As these effects diminish, sports performance will naturally reach equilibrium. Performance will only change in the future if an intervention takes place: this could be the emergence of a new athlete population, a new technology or a rule change.

5. Conclusions

The top-25 performances in 8 men's and 5 women's field events showed large increases in performance after the 2nd World War. A performance improvement index, set to 100% in 1948, was used to compare performances across events and showed that performance increased to 140.9% in throwing events and 125.8% in jumping events. Modelling the performance improvement index using a superposition of functions was able to quantify the underlying improvements and transient changes due to interventions such as rule changes, new technologies and performance-enhancing drugs.

It was shown that while technologies such as new javelins or vaulting poles could improve the performance improvement index, rules to limit their capabilities could reduce the performance improvement index by a similar or larger amount. Drugs testing and the formation of WADA were associated with reductions in the index for field events implying that drugs were in use by a large proportion of the top-25 prior to the introduction of these interventions. The drugs effect was consistently larger for women than men.

In a similar way to track events, field events were shown to have reached 99.9% of their predicted limit with all reaching it within the next 25 years. It was concluded that performance will only change in the future if an intervention takes place: this could be the emergence of a new athlete population, a new technology or a rule change.

References

Balmer, N., Pleasence, P. & Nevill, A. (2012), Evolution and revolution: Gauging the impact of technological and technical innovation on Olympic performance. Journal of Sports Sciences, 30(11) 1075-1083.

Berthelot, G., Tafflet, M., El Helou, N.,Len, S., Escolano, S., Guillaume, M., Nassif, H., Tolaïni, J., Thibault, V., Desgorces, F-D., Hermine, O., Toussaint, J-F., (2010). Athlete atypicity on the edge of human achievement: performances stagnate after the last peak, in 1988. Plos One 5(1), e8800.

Blest, D. C. (1996). Lower bounds for athletic performance. The Statistician, 45 (2), pp. 243-253.

Burgess, S. C. (1996). The design optimisation of poles for pole vaulting. In S. Haake (Ed.), The Engineering of Sport, Balkema, Rotterdam. pp 83–90.

Denny, M. W. (2008). Limits to running speed in dogs, horses and humans. Journal of Experimental Biology, 211, pp. 3836-3849.

Ernst, S., and Simon, P., (2013) A quantitative approach for assessing significant improvements in elite sprint performance: Has IGF-1 entered the arena? Drug Testing and Analysis. 5(6) 384-389.

Foster L (2012) The effect of technology on elite sport. PhD Thesis, Sheffield Hallam University, pp 343

Foster, L., James, D. & Haake, S. J. (2010). Understanding the influence of population size on athletic performance. Procedia Engineering, 2 (2), pp. 3183-3189.

Foster L., James, D. & Haake, S. (2011). The influence of the Olympic Games on athletic performance. Proceedings of the 3rd International conference on Mathematics in Sport.

Haake, S. J. (2009). The impact of technology on sporting performance in Olympic sports. Journal of Sport Sciences, 27(13), pp1421-1431.

Haake, S. J., James, D. & Foster, L. (2013). An improvement index to quantify the evolution of performance in running. Journal of Sport Sciences, (online).

IAAF. (2011). International Association of Athletics Federations - Home page [online]. Last accessed on 20/02/ 2011 at: http://www.iaaf.org/.

Lippi, G., Banfi, G., Favaloro, E.J., Rittweger, J. & Maffulli, N. (2008). Updates on improvement of human athletic performance: focus on world records in athletics. British Medical Bulletin, (87), 7-15.

Moré J. J. (1978). The Levenberg-Marquardt algorithm: implementation and theory, Lecture notes in mathematics, 630 (1978), pp. 105 – 116.

Norton, K., & Olds, T. (2001). Morphological evolution of ath- letes over the 20th century. Sports Medicine, 31(11), 763–783. Rabinovich,

Nevill, A. M. & Whyte, G. (2005). Are there limits to running World Records? Medicine & Science in Sports & Exercise, 37 (10), pp. 1785-1788.

Rabinovich, M. (2013). Track and field statistics. Retrieved from http://trackfield.brinkster.net

Ratkowsky, D. A. (1983). Nonlinear regression modelling: A unified practical approach. New York, NY: Marcel Dekker.

Seber, G. A. F. & Wild, C. J. (2003). Nonlinear regression. Hoboken, NJ: Wiley-Interscience.

Yang, M. C. K. (1975). On the distribution of the inter-record times in an increasing population. *Journal of Applied Probability*, 12 (1), pp. 148-154.

Table 1. Interventions and model parameters for the performance improvement index in men's and women's field events using a baseline of 1948. Interventions were introduced in the order shown.

Table 2. Components of the performance improvement index for the men's and women's field with data for track events from Haake *et al*. (2013).

Figure 1. Mean of the top-25 performances in 8 men's and 5 women's field events from 1890 to 2012.

Figure 2. The performance improvement index for 8 men's and 5 women's field events from 1948 to 2012 where 1948=100%. Also shown are the best-fit models using equation 10 and the parameters in Table 1. (Note: the scales are maximised for each event to allow detail to be seen).

Figure 3. The components of the performance improvement index for the men's high jump.

Figure 4. The change in performance improvement index between 1948 and 2012 due to positive and negative influences. The total length of each bar indicates the underlying secular change between 1948 and 2012 given by (3).

Table 1.

Step	Intervention	Model parameters (error)	R^2	R^2_{adj}	ΔR^2_{adj}	MSE	Date at 0.999L
Long jump men (5-parameter model)							
	Global improvement	L=1.1066 (0.011); a_1=0.193 (0.073); a_2=4.99 (1.35)	0.9572	0.9535	0.9335	0.000053	2022
1	Formation of WADA (2000)	c_1=-0.0167 (0.0078)			0.0149		
2	Olympic Games oscillation	A_2=0.0035 (0.0025)			0.0052		
Triple jump men (6-parameter model)							
	Global improvement	L=1.1698 (0.013); a_1=0.069 (0.025); a_2=8.91 (2.33)	0.9831	0.9813	0.9796	0.000044	2010
1	Formation of WADA (2000)	c_1=-0.0059 (0.0081)			0.0006		
2	Compulsory random drug testing (1989)	c_2=-0.0068 (0.0065)			0.0007		
3	Olympic Games oscillation	A_3=0.0018 (0.0023)			0.0005		
High jump men (7-parameter model)							
	Global improvement	L=1.1810 (0.031); a_1=0.127 (0.036); a_2=5.64 (1.27)	0.9890	0.9875	0.9726	0.000040	2030
1	Uptake of Fosbury flop (1968-1976)	Δc_1= 0.002184 yr^{-1} (0.001264)			0.0094		
2	Formation of WADA (2000)	c_2=-0.0187 (0.0086)			0.0038		
3	Compulsory random drug testing (1989)	c_3=-0.0107 (0.0089)			0.0015		
4	Olympic Games oscillation	A_4=0.0019(0.0022)			0.0002		
Pole vault men (7-parameter model)							
	Global improvement	L=1.3435 (0.044); a_1=0.027 (0.020); a_2=11.40 (5.3)	0.9927	0.9917	0.9857	0.00014	2026
1	Formation of WADA (2000)	c_1=-0.0356 (0.0156)			0.0025		
2	Uptake of composite poles (1956-1972)	Δc_2= 0.004966 yr^{-1} (0.002256)			0.0029		
3	Olympic Games oscillation	A_3=0.0039 (0.0041)			0.0006		
4	Compulsory random drug testing (1989)	c_4= 0.0111 (0.0157)			0.0000		
Shot put men (5-parameter model)							
	Global improvement	L=1.3463 (0.018); a_1=0.006 (0.002); a_2=31.6 (8.29)	0.9836	0.9822	0.9685	0.000178	2003
1	Compulsory random drug testing (1989)	c_1=-0.0515 (0.0164)			0.0013		
2	Olympic Games oscillation	A_2=0.0041 (0.0046)			0.0006		
Hammer men (6-parameter model)							
	Global improvement	L=1.5141 (0.031); a_1=0.0021 (0.0007); a_2=48.6 (11.5)	0.9890	0.9878	0.9846	0.000279	2007
1	Compulsory random drug testing (1989)	c_1=-0.0338 (0.0224)			0.0021		
2	Formation of WADA (2000)	c_2= -0.0161 (0.0165)			0.0005		
3	Olympic Games oscillation	A_3= 0.0056 (0.0058)			0.0005		
Discus men (5-parameter model)							
	Global improvement	L=1.3690 (0.017); a_1=0.005(0.001); a_2=36.5 (8.32)	0.9880	0.9870	0.9734	0.000154	2003
1	Compulsory random drug testing (1989)	c_1=-0.0535 (0.0154)			0.0128		
2	Olympic Games oscillation	A_2=0.0045 (0.0043)			0.0007		
Javelin men (6-parameter model)							
	Global improvement	L=1.3264 (0.036); a_1=0.040 (0.016); a_2=9.53 (2.91)	0.9814	0.9795	0.8836	0.000116	2027
1	Specification change (1986)	c_1=-0.1052 (0.0141)			0.0769		
2	Uptake of hollow javelins (1953 - 1956)	Δc_2= 0.019358 yr^{-1} (0.005219)			0.0130		
3	Formation of WADA (2000)	c_3=-0.0267 (0.0154)			0.0059		

Long jump women (6-parameter model)							
	Global improvement	L=1.2946 (0.035); a_1=0.066 (0.018); a_2=7.25 (1.58)	0.9827	0.9809	0.9609	0.000091	2034
1	Compulsory random drug testing (1989)	c_1=-0.0230(0.0121)			0.0063		
2	Formation of WADA (2000)	c_2=-0.0339 (0.0121)			0.0119		
3	Olympic Games oscillation	A_1=0.0043 (0.0033)			0.0018		
High jump women (6-parameter model)							
	Global improvement	L=1.2219 (0.029); a_1=0.060 (0.015); a_2=8.61 (1.73)	0.9947	0.9939	0.9765	0.000044	2020
1	Uptake of Fosbury flop (1968 - 1976)	Δc_1= 0.005994 yr^{-1} (0.001397)			0.0153		
2	Olympic Games oscillation	A_2= 0.0035 (0.0023)			0.0010		
3	Compulsory random drug testing (1989)	c_3= -0.0150 (0.0097)			0.0007		
4	Formation of WADA (2000)	c_4= -0.0103 (0.0085)			0.0005		
Shot put women (5-parameter model)							
	Global improvement	L=1.6051 (0.04); a_1=0.0003 (0.0001); a_2=144.26(46.0)	0.9792	0.9770	0.9596	0.000668	2001
1	Compulsory random drug testing (1989)	c_1= -0.1137 (0.0333)			0.0239		
2	Formation of WADA (2000)	c_2= -0.0380 (0.0226)			0.0027		
3	Olympic Games oscillation	A_3= 0.0080 (0.0090)			0.0013		
Discus women (5-parameter model)							
	Global improvement	L=1.6541 (0.029); a_1=0.0002 (0.00008); a_2=166.3 (39.9)	0.9887	0.9876	0.9596	0.000402	2001
1	Compulsory random drug testing (1989)	c_1= -0.1030(0.0259)			0.0239		
2	Formation of WADA (2000)	c_2= -0.0343 (0.0175)			0.0027		
3	Olympic Games oscillation	A_3= 0.0095 (0.0070)			0.0013		
Javelin women (5-parameter model)							
	Global improvement	L=1.6332 (0.05); a_1=0.002 (0.0006); a_2=42.6 (8.8)	0.9892	0.9877	0.9492	0.000308	2019
1	Specification change (1999)	c_1=-0.0900 (0.0252)			0.0219		
2	Compulsory random drug testing (1989)	c_2=-0.0674 (0.0266)			0.0107		
3	Uptake of hollow javelins (1953 - 1956)	Δc_3= 0.014278 yr^{-1} (0.007507)			0.0021		
4	Tail roughness rule (1992)	c_4= -0.0525 (0.0261)			0.0030		
5	Olympic Games oscillation	A_5= 0.0069 (0.0061)			0.0009		

Table 2.

	Long jump	Triple jump	High jump	Pole vault	Shot put	Hammer	Discus	Javelin	Jumping	Throwing	**All field**	100 m	200 m	400 m	800 m	1,500 m	5,000 m	10,000 m	Marathon	Sprint	Middle	Long	**All track**	**All**
i) Performance improvement index 2012 in % (equation 10): Field n=13; Track n=11																								
	Field events											Track events												
Men	109.1	115.8	116.4	137.1	129.9	146.9	132.0	124.4	119.6	133.3	**126.5**	110.5	112.1	111.9	113.1	117.6	123.6	121.2	146.7	111.5	115.4	130.5	**119.6**	**120.0**
Women	122.3		129.3		146.1		152.6	146.7	125.8	148.5	**139.4**	120.7	126.5		143.3					123.6	143.3		**130.2**	**132.4**
All	115.7	115.8	122.8	137.1	138.0	146.9	142.3	135.6	122.7	140.9	**132.9**	115.6	119.3	111.9	128.2	117.6	123.6	121.2	146.7	117.6	129.3	130.5	**124.9**	**124.6**
(ii) Exponential rise 1948-2012 in % (equation 6): Field n=13; Track n=11																								
Men	110.4	116.9	117.4	133.5	134.6	151.4	136.9	131.8	119.5	138.7	**129.1**	116.2	116.4	115.7	111.6	115.4	119.2	122.7	139.9	116.1	113.5	127.3	**119.6**	**120.2**
Women	128.2		121.9		160.5		165.4	162.7	125.0	162.9	**147.7**	127.3	131.3		146.7					129.3	146.7		**135.1**	**137.7**
All	119.3	116.9	119.6	133.5	147.6	151.4	151.2	147.3	122.3	150.8	**138.4**	121.8	123.9	115.7	129.1	115.4	119.2	122.7	139.9	122.7	130.1	127.3	**127.4**	**125.7**
(iii) Contribution from compulsory random drugs testing (1989) in %: Field n=11; Track n=11																								
Men		-0.7	-1.1	-1.1	-5.1	-3.4	-5.4		-1.0	-4.6	**-2.8**	-1.3	-0.9	-1.7	-1.4	-1.8	-0.8	-1.5	-2.5	-1.3	-1.6	-1.6	**-1.5**	**-1.6**
Women	-3.0		-1.5		-11.4		-10.3	-6.7	-2.2	-9.5	**-6.6**	-2.3	-3.0		-3.9					-2.6	-3.9		**-3.1**	**-3.6**
All	-3.0	-0.7	-1.3	-1.1	-8.3	-3.4	-7.8	-6.7	-1.6	-7.0	**-4.7**	-1.8	-1.9	-1.7	-2.7	-1.8	-0.8	-1.5	-2.5	-2.0	-2.8	-1.6	**-2.3**	**-2.1**
(iv) Contribution from formation of WADA (2000) in %: Field n=10; Track n=7																								
Men	-1.7	-0.6	-1.9	-3.6		-1.6		-2.7	-1.9	-2.1	**-2.0**	-0.6	-0.7	-2.0	-0.9	-0.5				-1.1	-0.7		**-0.9**	**-1.0**
Women	-3.4		-1.0		-3.8		-3.4		-2.2	-3.6	**-2.9**	-2.1	-2.5							-2.3			**-2.3**	**-2.4**
All	-2.5	-0.6	-1.4	-3.6	-3.8	-1.6	-3.4	-2.7	-2.1	-2.9	**-2.5**	-1.3	-1.6	-2.0	-0.9	-0.5				-1.7	-0.7		**-1.6**	**-1.4**
(v) Contribution from Olympic oscillation in %: Field n=12; Track n=11																								
Men	0.4	0.2	0.2	0.4	0.4	0.6	0.4		0.3	0.5	**0.4**	0.3	0.4	0.4	0.2	0.2	0.2	0.3	0.3	0.4	0.2	0.3	**0.3**	**0.3**
Women	0.4		0.4		0.8		0.9	0.7	0.4	0.8	**0.6**	0.7	0.7		0.5					0.7	0.5		**0.6**	**0.6**
All	0.4	0.2	0.3	0.4	0.6	0.6	0.7	0.7	0.3	0.6	**0.5**	0.5	0.5	0.4	0.3	0.2	0.2	0.3	0.3	0.5	0.3	0.3	**0.5**	**0.4**

(vi) Residual global effect in %: Field n=13; Track n=11

Men	8.3	15.4	12.5	20.5	29.1	45.8	31.1	12.8	14.2	29.7	**22.0**	8.7	11.2	12.1	5.5	8.6	12.9	16.0	27.8	10.7	7.0	18.9	**12.9**	**13.4**
Women	21.4		14.2		44.6		50.7	36.8	17.8	44.0	**33.5**	19.3	25.2		42.3					22.2	42.3		**28.9**	**30.5**
All	14.9	15.4	13.4	20.5	36.8	45.8	40.9	24.8	16.0	36.9	**27.7**	14.0	18.2	12.1	23.9	8.6	12.9	16.0	27.8	16.5	24.6	18.9	**20.9**	**18.6**

(vii) Contribution from technology and population influx in %

			Field events n=8								Track events n=10 (Haake *et al.* 2013)										
	Intervention		Long jump	Triple jump	High jump	Pole vault	Shot put	Hammer	Discus	Javelin	100 m	200 m	400 m	800 m	1,500 m	5,000 m	10,000 m	Marathon	Mean
Positive influences	Fosbury Flop uptake (1968-1976)	Men			1.7														
	Fosbury Flop uptake (1968-1975)	Women			4.8														
	Composite poles uptake (1956-1972)	Men				7.9													4.9
	Hollow javelin uptake (1953-1956)	Men								5.8									
	Hollow javelin uptake (1953-1956)	Women								4.3									
	Influx of African runners *	Men												3.6	4.4	5.3	4.9	9.3	4.7
	Usain Bolt effect (2008)	Men									0.6								
	All																		**4.8**
Negative influences	COM rule change (1986)	Men								-10.5									
	COM rule change (1999)	Women								-9.0									-8.3
	Tail roughness (1992)	Women								-5.3									
	Fully automated timing (1975)	Men									-4.6	-3.2	-0.5						-2.8
	Fully automated timing (1975)	Women									-2.9								
	All																		**-5.1**

* 800 m and 1,500m 1980-2000; 5,000 m 1980-2003; 10,000 m 1980-2007; Marathon 1980-2009.

Figure 1.

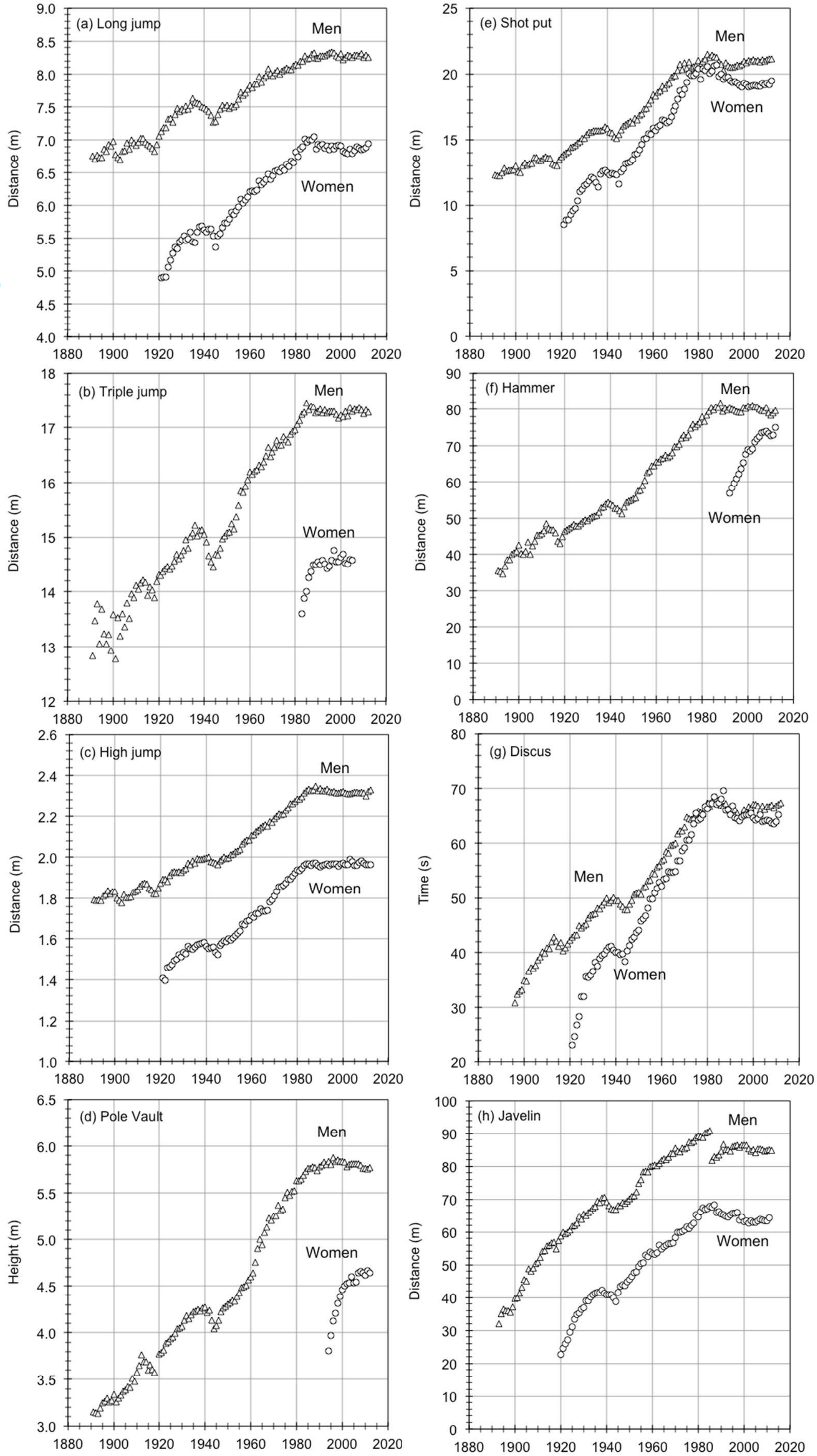

(a) Long jump
Men
Women
Distance (m)
(e) Shot put
Men
Women
Distance (m)
(b) Triple jump
Men
Women
Distance (m)
(f) Hammer
Men
Women
Distance (m)
(c) High jump
Men
Women
Distance (m)
(g) Discus
Men
Women
Time (s)
(d) Pole Vault
Men
Women
Height (m)
(h) Javelin
Men
Women
Distance (m)
1880
1900
1920
1940
1960
1980
2000
2020

Figure 2.

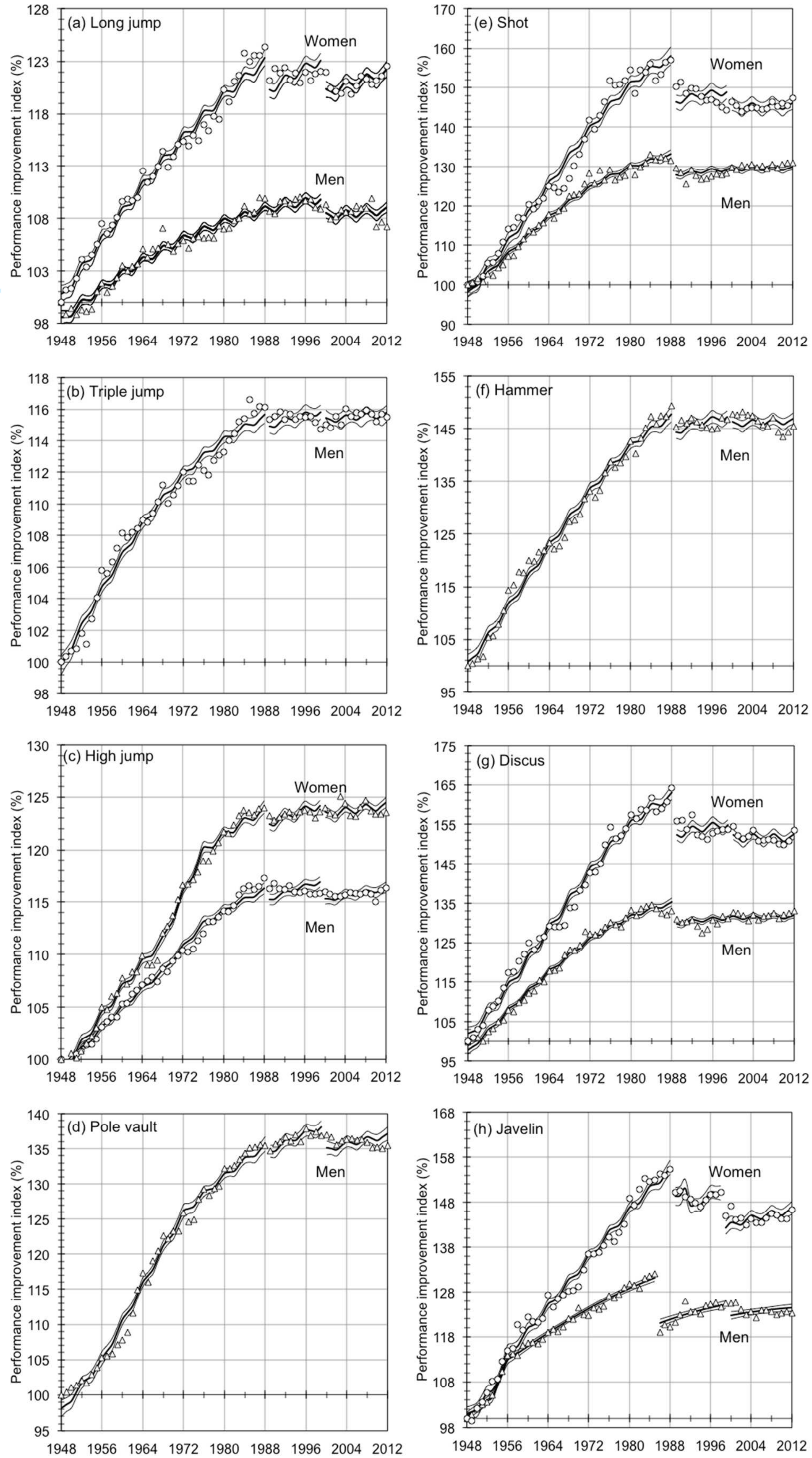

(a) Long jump
Women
Men
(b) Triple jump
Men
(c) High jump
Women
Men
(d) Pole vault
Men
(e) Shot
Women
Men
(f) Hammer
Men
(g) Discus
Women
Men
(h) Javelin
Women
Men
Performance improvement index (%)
1948 1956 1964 1972 1980 1988 1996 2004 2012

Figure 3.

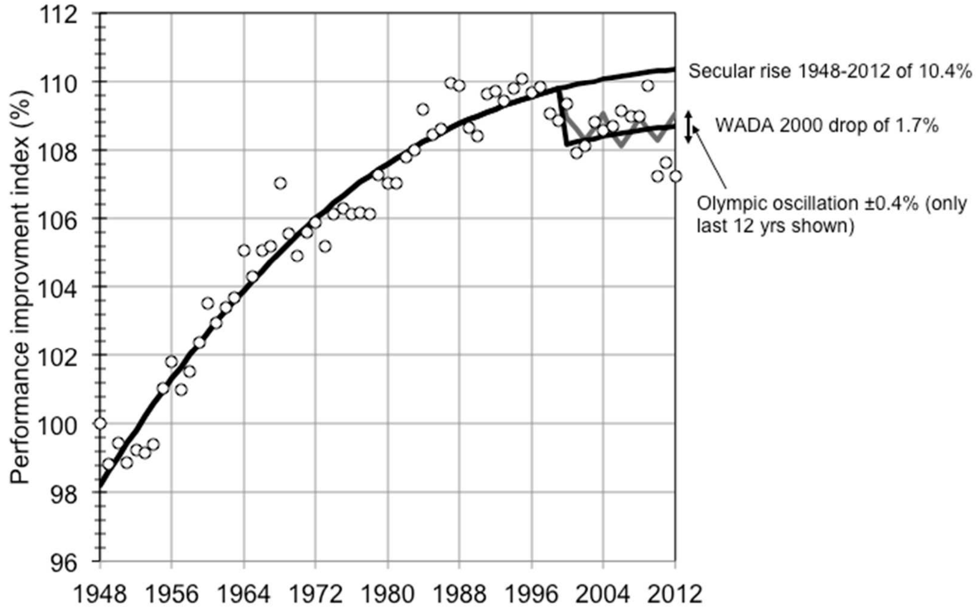
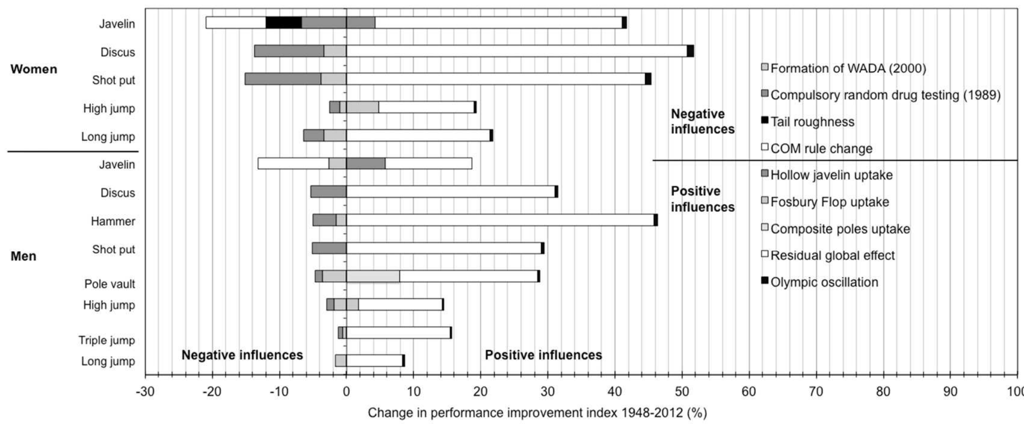

Performance improvment index (%)
112
110
108
106
104
102
100
98
96
1948
1956
1964
1972
1980
1988
1996
2004
2012
Secular rise 1948-2012 of 10.4%
WADA 2000 drop of 1.7%
Olympic oscillation ±0.4% (only last 12 yrs shown)

Figure 4.


Javelin
Discus
Women
Shot put
High jump
Long jump
Javelin
Discus
Hammer
Men
Shot put
Pole vault
High jump
Triple jump
Long jump
Negative influences
Positive influences
Formation of WADA (2000)
Compulsory random drug testing (1989)
Tail roughness
COM rule change
Negative influences
Positive influences
Hollow javelin uptake
Fosbury Flop uptake
Composite poles uptake
Residual global effect
Olympic oscillation
-30 -20 -10 0 10 20 30 40 50 60 70 80 90 100
Change in performance improvement index 1948-2012 (%)