DEVELOPING METHODS TO ASSESS THE RELATIONSHIP BETWEEN ERGOMETER AND ON-WATER ROWING PERFORMANCE FROM INDEPENDENT DATASETS

Sam Jones[1], Chris Bailey[2], Dave Thomas[2], Mark G. E. White[1], Paul Rees[3], Huw D. Summers[3], Neil E. Bezodis[1]

Applied Sports, Technology, Exercise and Medicine Research Centre, Swansea University, Swansea, UK[1]
English Institute of Sport, Manchester, UK[2]
Department of Biomedical Engineering, Swansea University, Swansea, UK[3]

Rowing ergometers are often used by internationally competitive athletes alongside on-water rowing. This study proposes methods to develop a generalisable relationship between maximal effort 2000 m ergometer and on-water rowing performance using independent datasets. Ergometer times for 2000 m tests (*n* = 153) and 2000 m on-water times from international races (*n* = 139) were collated. Percentiles from the raw data and fitted probability density functions were mapped to develop a generalisable performance relationship. Bootstrapping was utilised to estimate the uncertainty in the percentile mappings. When built on a larger sample of athletes, this approach could be useful to identify athletes who under or overperform on water compared to ergometers, and this could provide valuable context for future biomechanical investigations of rowing technique.

KEYWORDS: bootstrap, distribution, single scull, technique

INTRODUCTION: Rowing performance is measured by the ability to cover a set distance in the fastest time possible. Rowing practice often takes place in boats on water, but rowing ergometers are also commonly used for training, selection, and physiological performance testing and monitoring (Smith & Hopkins, 2012). Despite the common use of both on-water and ergometer rowing in practice, races occur on water and the relationship in performance between the two modes of rowing is not clear; some athletes may perform better or worse when rowing on water than might be expected from their respective ergometer performance.
Measures of ergometer and on-water performance do appear to be positively correlated. Significant positive correlations have been observed between 2000 m ergometer and on-water times for 15 male single scullers at a national team selection camp (*r* = 0.90; Barrett & Manning, 2004) and for 10 experienced male single scullers (*r* = 0.72; Jürimäe et al., 2000). However, a weaker non-significant correlation (*r* = 0.12) was reported for 19 male athletes during a national team training camp (McNeely, 2011). Whilst it is likely that there is a positive relationship between 2000 m ergometer and on-water rowing performance, the strength of the relationship differs between studies and it is not a perfect correlation. This confirms that some athletes perform better or worse when rowing on water than might be expected based solely on their ergometer performance. Furthermore, all studies analysed small datasets meaning that athletes with outlying performances would have a greater influence on the relationship.
Having a generalisable performance relationship between the two modes of rowing – one which any athlete from a wider population of internationally competitive male single scullers could be evaluated against – would enable the identification and quantification of these potentially under or overperforming athletes. This metric would then provide valuable context for subsequent biomechanical investigations of rowing technique. Building a performance relationship would ideally be reliant on two dependent datasets with sufficiently large size where each measure within each dataset is independent. However, acquiring such datasets is hard to achieve from an internationally competitive male single scull population. Therefore, independently-gathered datasets are likely required given the challenges of collecting both datasets from a single cohort and this study will develop and propose potential methods for using independent datasets to build a generalisable relationship between 2000 m ergometer and on-water rowing performance. Whilst these independent datasets should still each comprise independent measures, this study will use multiple samples from individuals for the purposes of illustrating and evaluating the proposed methods, which can consequently then

be applied to larger independent datasets when available. The aim of the current study was to develop an approach for determining a generalisable relationship between independent datasets of 2000 m ergometer and on-water rowing performance.

METHODS: Times for maximal effort 2000 m ergometer rowing were extracted from the British Rowing database for internationally competitive athletes who had competed or trained in men's single sculling from February 2015 and had left the programme by December 2021. Men's on-water single scull race times from the same time period were collected from worldrowing.com. These consisted of race times achieved by all athletes in A and B finals during international competitions (Olympics, World Rowing Championships, World Rowing Cup IIIs). For this study, repeated measures were included within each dataset to ensure sufficiently large datasets were obtained to enable a more complete evaluation of the different approaches tested. The ergometer dataset consisted of 153 performance times from 12 athletes (range = 2-20 per athlete). The on-water dataset consisted of 139 performance times from 55 athletes (range = 1-10 per athlete). Although the datasets were collected independently meaning that most of the athletes in the ergometer dataset were not in the on-water dataset, some athletes were likely included in both datasets due to the anonymised nature of the ergometer data.
Each percentile from the on-water dataset was mapped one-to-one in ascending order (i.e. bijectively) to the corresponding percentile from the ergometer dataset to provide a generalisable performance relationship. A two-parameter probability density function (PDF; right-skewed Gumbel) was fitted to each of the raw ergometer and on-water datasets to provide smooth estimations of the collected distributions. Percentiles derived from these PDFs were then also plotted against each other in the same manner as those from the raw data.
Bootstrapping was utilised to estimate the uncertainty in the two methods of percentile mapping. The bootstrap involved randomly sampling with replacement from the raw ergometer and on-water datasets to produce 1000 bootstrap samples of the same sizes as the raw datasets. For method one, percentiles from these bootstrap samples were bijectively mapped. For method two, PDFs (right-skewed Gumbel) were fitted to each bootstrap sample and the percentiles from these PDFs were bijectively mapped. The median and interquartile range (IQR) were determined for each percentile across all bootstrap samples. All data analysis was completed using Python 3.10.8 (Python Software Foundation, python.org).
To illustrate how the proposed methods could be used to evaluate individual athletes (from the same wider population) against the developed relationship, four additional athletes with both a 2000 m ergometer and on-water time were overlaid on a visual representation of the relationship (labelled A-D on Figure 1c, and H is a hypothetical athlete).

RESULTS: Distributions for the ergometer and on-water datasets are shown in Figure 1a and 1b, respectively. The location and scale parameters were 355.6 and 4.9, respectively, for the PDF fitted to the ergometer distribution and 411.4 and 10.0, respectively, for the on-water distribution. The percentile mapping lines from the raw distributions (solid black line) and the fitted PDFs (dashed black line) are shown in Figure 1c. The percentile mappings from the bootstrap samples are shown in Figure 2a and 2b, respectively. The median and IQR of each percentile across the 1000 samples were also plotted (dashed black lines in Figure 2a and 2b).

DISCUSSION: This study investigated how a generalisable relationship between 2000 m ergometer and on-water rowing performance could be developed using independent datasets from the same wider population (i.e. internationally competitive male single scullers). Knowing the properties of such a relationship (i.e. the two percentile mapping lines in Figure 1c) would enable the evaluation of athletes with both ergometer and on-water performances, such as athletes A-D in Figure 1c. It has been suggested that some athletes perform better or worse when rowing on water than might be expected based on their ergometer performance (McNeely, 2011), and whilst coaches often qualitatively assess this, it can be challenging to quantify. Such a metric can be used as a dependent variable in biomechanical investigations of on-water rowing technique, which may be contributing to any under or overperformance.

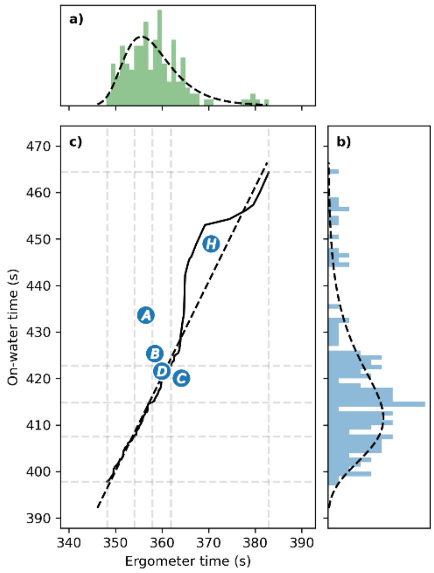

Figure 1: Distributions of 2000 m a) ergometer and b) on-water times. The dashed black lines in a) and b) illustrate fitted probability density functions. Percentile mappings from the raw data (solid black line) and fitted probability density functions (dashed black line) are shown in c), along with 25 percentile increments (grey dashed lines) and 2000 m ergometer and on-water times for four athletes (labelled A-D) and one hypothetical athlete (labelled H).

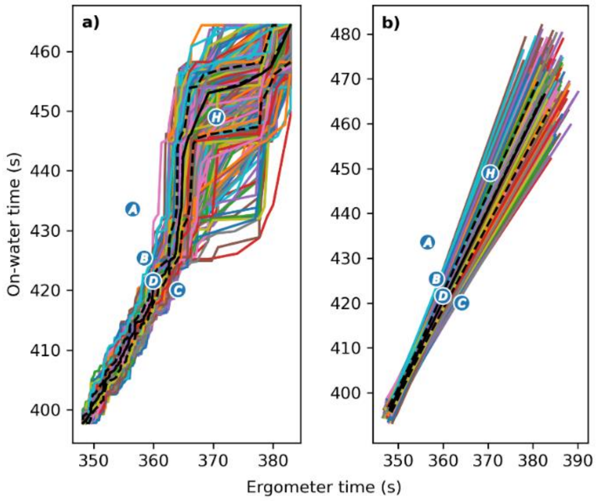

Figure 2: Uncertainty in the two methods of percentile mapping as illustrated by the bootstrapping methods. a) Percentiles for each ergometer and on-water bootstrap sample were bijectively mapped. b) A probability density function was fitted to each ergometer and on-water bootstrap sample and the percentiles from these fitted probability density functions were then bijectively mapped. The solid black lines illustrate the raw percentile mappings (as seen in Figure 1c) and the dashed black lines indicate the median and interquartile range across the bootstrap samples. Ergometer and on-water 2000 m times for four athletes (labelled A-D) and one hypothetical athlete (labelled H) are overlaid.

Ideally, both datasets would come from the same sample of athletes and each athlete would contribute a single performance time. However, as is evident from the lack of existing studies investigating the relationship between 2000 m ergometer and on-water rowing, it is challenging to obtain this information from a sufficiently large sample of internationally competitive athletes. This study therefore develops proposed methods to obtain a more general performance relationship from larger independent datasets. It should be noted, however, that repeated measures have been included in each of the current datasets to provide a sufficient sample size to enable evaluation of the different approaches. Future work to develop a relationship using the proposed methods should only be implemented using fully independent datasets when larger samples are available. Regardless, the current study proposes an approach for managing the limitation of using independent datasets to develop a generalisable performance relationship. Evidence suggests there is likely a positive relationship between ergometer and on-water performance (Barrett & Manning, 2004; Jürimäe et al., 2000) and bootstrapping was used to give an indication of the uncertainty in the proposed percentile mappings. The implications of the bijective percentile mappings should be explored further, and different mappings built under different assumptions could also be explored.

Whilst the use of independent datasets is a fundamental limitation which this approach endeavours to address, if the samples of athletes in each dataset are sufficiently large with no repeated measures, and if they are drawn from the same wider population (internationally competitive male single scullers), then confidence can be placed in the developed relationship. The datasets included in the current study are still relatively small and not without bias, but the aim of this study was to develop proposed methods for working with independent datasets. Future work to increase these samples whilst ensuring they are fully independent would make the performance relationship more generalisable and applicable to the population of internationally competitive male single scullers.

The under or overperformance when rowing on water for any internationally competitive male single sculler could be identified by their vertical residual from a developed generalisable performance relationship line when their actual times are overlaid. For example, in Figure 1c athlete A is above the percentile lines and so would be deemed as an ‘underperformer’. On the other hand, athlete C is below the percentile lines and would therefore be deemed as an ‘overperformer’. Caution must be exerted though in overinterpreting athletes who lie close to any modelled relationship lines, e.g. athlete D. Subsequent biomechanical investigations into an athlete’s rowing technique may help uncover the underlying causes of their under or overperformance and therefore identify areas for improvement. The hypothetical athlete (athlete H) indicates how using different methods to develop a generalisable performance relationship would result in different under or overperformance categorisation (Figure 1c). In this region of the figure there would be less certainty in the under or overperformance categorisation (as illustrated by the bootstrapping methods and IQR lines in Figure 2a and 2b) due to the lower density of ergometer and on-water 2000 m times (shown in Figure 1a and 1b). The lack of density is a feature that could be expected in such performance data since the data are from internationally competitive athletes who all perform to a similar high standard, and so there are fewer times recorded towards the slower extreme. Fitting PDFs to the raw datasets appears to be preferable for developing a generalisable performance relationship since PDFs can help overcome the uncertainty where there is a lack of density and reduce the effects of random variation in the original distributions. A relationship developed using PDFs is consequently less prone to the effects of peculiarities in the original datasets, as is evident from the narrower spread of percentile mappings and IQR from the bootstrapping (Figure 2b).

CONCLUSION: Methods were developed that could be used to obtain a generalisable performance relationship between independent datasets of 2000 m ergometer and on-water rowing times from the population of internationally competitive male single scullers. Obtaining such a performance relationship from larger truly independent datasets in the future would enable the identification of athletes who perform better or worse on water than might be expected from their respective ergometer performance. This could subsequently provide context for biomechanical investigations of rowing technique.

REFERENCES

Barrett, R. S., & Manning, J. M. (2004). Relationships between rigging set-up, anthropometry, physical capacity, rowing kinematics and rowing performance. *Sports Biomechanics*, *3*(2), 221–235.

Jürimäe, J., Mäestu, J., Jürimäe, T., & Pihl, E. (2000). Prediction of rowing performance on single sculls from metabolic and anthropometric variables. *Journal of Human Movement Studies*, *38*, 123–136.

McNeely, E. (2011). Rowing ergometer physiological tests do not predict on-water performance. *The Sport Journal*, *14*, 12–15.

Smith, T. B., & Hopkins, W. G. (2012). Measures of rowing performance. *Sports Medicine*, *42*(4), 343–358.