

Моделирование результатов легкоатлетов на основе электронного мониторинга пульсовых показателей

А.И. Епанчинцева, М.А. Бакаев

Аннотация— Возможность предсказывать способности и будущие результаты спортсменов важна для эффективного распределения усилий тренеров, времени использования спортивного оборудования, и других ограниченных ресурсов, доступных в рамках проектов по развитию спорта. Применение для решения этой задачи популярных сейчас методов машинного обучения ограничено требованием больших объемов данных и наличия статистики предыдущих тренировок и выступлений спортсменов. В нашей статье мы описываем пилотное исследование, проведенное с 14 действующими бегунами на средние дистанции, и предлагаем регрессионную модель для предсказания результатов, которая объясняет 89% их дисперсии. Модель основана на пульсовых показателях в различных видах тренировок – т.е. данных, которые могут быть относительно легко собраны современными электронными сенсорами или мобильными устройствами. Для времени бега на дистанции в 1 км значимыми оказались факторы пульса на восстановительном кроссе, на рабочем кроссе и максимальный пульс спортсмена. Мы полагаем, что наши результаты могут быть полезны как легкоатлетам и тренерам, так и лицам, принимающим решения в сфере спортивного менеджмента – в частности, для составления оптимальных графиков тренировок и отборов в сборные команды страны.

Ключевые слова— индикаторы пульса, медицинские информационные системы, спортивный менеджмент, спортивные показатели.

I. ВВЕДЕНИЕ

Современный международный спорт – это соревнование быстрых управленческих решений, высоких технологий и сильных инвестиций. В частности, для легкой атлетики – одного из самых древних и престижных видов спорта – большое значение имеет эффективность отбора кандидатов в спортсмены для дальнейшего их развития. Например, в беге, требующем выносливости, т.е. на средние, длинные и сверхдлинные дистанции,

крайне желательно заранее предсказать, каких результатов в принципе способен достичь атлет. Это позволяет более эффективно распределять ограниченные ресурсы, в том числе время работы тренеров, а также формировать сборные команды наилучшим образом. Качественное моделирование на основе спортивных данных и предсказание результатов имеет особую важность для российского спорта [1], поскольку наши атлеты в последнее время отстают как в мужских, так и в женских видах.

Вообще, в области спортивных показателей в настоящее время широко применяются технологии машинного обучения, а наиболее популярной задачей является предсказание победителей [2], в особенности футбольных матчей (см., например, [3]). Это легко объяснимо, поскольку в футболе накоплено, пожалуй, наибольшее количество статистики, а экономический эффект от правильного прогнозирования результатов может быть весьма значительным. Недавний обзор более 100 публикаций в этой области [4] констатирует, что в большинстве из них проблема прогнозирования ставится как классификация, что по понятным причинам не является адекватным подходом при прогнозировании индивидуальных результатов. В этом же обзоре авторы отмечают, что наиболее широко используемый метод – это нейронные сети на основе признаков. Как известно, они требуют значительных объемов данных, однако в случаях, когда применяются методы глубокого обучения (как, например, в [5]), наборы данных нужны **еще в разы больше**. Такая «прожорливость» в отношении данных может быть отчасти удовлетворена за счет автоматических методов их сбора, например, из онлайн-источников [6]. Однако такой подход может быть применим только для уже давно выступающих команд и спортсменов, для которых накоплена статистика.

В то же время, развитие мобильных устройств и носимых сенсоров дало толчок развитию персональных рекомендательных систем [7]. Конечно же, исходя из своего назначения, они учитывают индивидуальные характеристики спортсмена, однако, как правило, они тоже требуют значительных объемов статистики и/или дневников тренировок. Между тем, как мы отмечали выше, для эффективной массовой организации спорта на основе информационных технологий весьма желательно иметь возможность предсказывать

Статья получена 10 октября 2022 г.

Епанчинцева Алина Игоревна, магистрант направления «Биотехнические системы и технологии» (alusik_e@mail.ru).

Бакаев Максим Александрович, кандидат технических наук, Новосибирский государственный технический университет, доцент, ORCID 0000-0002-1889-0692 (bakaev@corp.nstu.ru).

Исследование поддержано Российским фондом фундаментальных исследований, проект № 19-29-01017. Статья подготовлена по итогам выступления на Международной объединённой конференции «Интернет и современное общество» (IMS-2022).

возможности и будущие результаты атлетов **без предварительного вложения времени и сил** в тренировки на профессиональном уровне.

Примечательный подход, не требующий наличия дневников тренировок или истории выступлений, представлен в [8]. Авторы предсказывают результаты бегунов на длинные дистанции на основе характеристик личности, а конкретно, степени перфекционизма. И действительно, важность настойчивости в спорте было бы отрицать странно – в частности, от неё зависит тренировочная дисциплина, которая, в свою очередь, обуславливает более высокие результаты. Но нельзя не заметить, что характеристики личности могут, в свою очередь, приобретаться и развиваться в ходе собственно тренировок, поэтому использование их в качестве факторов для долгосрочных прогнозов кажется ненадежным. Представляется, что такие прогнозы имеют смысл скорее основываясь на физиологических особенностях, которые зачастую являются врожденными и менее подвержены изменениям.

При этом прогнозирование на основе биомеханических параметров является известной, но пока не решенной проблемой [9]. Работы в этой области всё больше нацеливаются на применение сложных методов, требующих дорогостоящего оборудования и аккуратного его применения. Представляется, что это не всегда возможно и не желательно в массовом спорте, в особенности в такой большой и неравномерно развитой стране, как Россия. Относительно недавно многообещающие модели были построены на основе данных с относительно простых электро-кардио сенсоров [10], но при этом целью авторов было прогнозирование (избегание) травм из-за перетренировок. Однако в целом мы полагаем, что показатели, основанные на сердцебиениях (пульсовые показатели), имеют хороший потенциал с точки зрения прогнозирования результатов спортсменов, и именно их мы более подробно рассматриваем в нашей работе. Результаты, представленные в данной статье, были предварительно доложены на международной конференции «Internet and Modern Society» (г. Санкт-Петербург) в июне 2022 г. [11].

В последующем разделе 2 нашей статьи мы описываем теоретический базис исследования, вкратце излагая особенности тренировочного процесса легкоатлетов, а также основанное на показателях пульса предсказывание их результатов. В разделе 3 мы анализируем данные, которые мы собрали у 14 действующих спортсменов, и строим модель линейной регрессии для времени бега на типичную дистанцию в 1 км. В наилучшей из моделей, 5 значимых пульсовых показателей в совокупности с демографическими данными **объясняют около 89% дисперсии результатов**. В Заключении мы формулируем и обсуждаем выводы из полученных нами результатов, а также предлагаем некоторые направления для дальнейших исследований.

II. ОБЗОР И МЕТОДЫ ИССЛЕДОВАНИЯ

A. Тренировочный процесс легкоатлетов

В рамках отражения теоретического базиса нашего исследования, здесь мы описываем тренировочный процесс действующей легкоатлетки (одна из авторов данной статьи), выступающей на дистанциях 800-1500 м. Для таких дистанций осуществляется специализированная тренировка, поскольку они требуют от спортсмена как скоростной, так и силовой выносливости. Необходимо отметить, что любой график тренировок основывается на ранее определенных характеристиках легкоатлета, но никогда не является окончательным. Оптимизация тренировочного процесса продолжается по ходу него, для достижения спортсменом наилучших результатов в нужный момент.

Самый важный момент в тренировке легкоатлета – это верный выбор нагрузки: имеет значение каждый километр дистанции и каждый день восстановления. Восстановление как минимум не менее важно, чем нагрузка, поскольку улучшение результатов спортсмена происходит именно во время его отдыха. Если баланс между двумя этими составляющими отсутствует, то сердце и сердечнососудистая система в целом могут не справиться с последующими объемами тренировок. Это одна из очень индивидуальных физиологических характеристик: одному спортсмену может потребоваться 2-5 дней отдыха в той же ситуации, когда другому требуется всего 1 день.

Рассмотрим задание тренера по ускоренному бегу на 1000 м, 8 раз с перерывами по 3 минуты. В последующий день рассматриваемая легкоатлетка будет иметь пульс 170 ударов в минуту во время даже обычного рабочего кросса. Она не получит улучшения результатов от такого тренировочного графика, в частности потому, что потребное восстановление не обеспечивается. Соответственно, её тренер на следующий день должен выбрать легкую тренировку, которая даст лишь 150-160 ударов в минуту во время кросса, и тогда после восстановления спортсменка может ожидать повышения результатов.

Еще один важный момент – это силовая тренировка, которая обычно осуществляется в тренажерном зале. Опять же, график тренировок определяется «обратной связью» от мышц и в значительной степени зависит от того, на какой дистанции специализируется легкоатлет. Так, бегуны на короткие дистанции (спринтеры) полагаются на физическую силу и должны стимулировать рост мускулов посредством подъема тяжелых весов в рамках малого количества подходов. Напротив, бегунам на длинные дистанции (стайерам) требуются мышечная выносливость, поэтому им следует заниматься с умеренным весом и делать короткие перерывы на отдых. Эффект от силовых тренировок длится 10-14 дней, поэтому необходимо тщательное планирование графика, чтобы избежать смазывания или наложения различных воздействий.

В качестве ещё одного примера специализированного приёма можно привести бег в гору, который может являться очень эффективным видом тренировки, особенно в случае совмещения с тренировкой выносливости. Конечно же, бег в гору улучшает силу ног, делает бег более эластичным, а шаг – длиннее. Но еще более важно то, что он эффективно повышает скорость бега на анаэробном пороге (АП) – точке, в которой сбалансированы выделение и разложение молочной кислоты в организме. Восстановление между забегами, которые обычно не длиннее 400 м, как правило осуществляется в ходе 2,5-минутных перерывов, препятствующих закислению. Бег в гору не вызывает значительного напряжения гормональной и вегетативной систем, но сердечно-сосудистые показатели по-прежнему остаются наиболее информативными показателями возможностей спортсмена.

В. Предсказание результатов спортсменов

В настоящее время тестирование с газоанализатором приобретает популярность как один из способов получения информации о пульсовых зонах спортсмена, АП и максимальном потреблении кислорода. Однако соответствующее оборудование является довольно сложным и дорогим, так что простые устройства для мониторинга частоты сердечных сокращений (ЧСС) по-прежнему широко распространены. Они регистрируют количество сердцебиений (обычно за минуту) и могут применяться как при отдыхе и восстановлении, так и при интенсивной тренировке.

Для целей повышения её эффективности выделяют несколько тренировочных этапов и зон. Первая – это разминка, когда пульс достигает 50-60% от максимального уровня. Следующая – анаэробная зона, подразумевающая активную работу лёгких по мере того, как организм начинает потреблять большие объемы кислорода. В этой зоне легкоатлеты осуществляют скоростные тренировки, и пульс поднимается до 80-90% от максимального. Наконец, дальше идёт пороговая зона, в которой как пульс, так и потребляемый кислород достигают 90-100% от максимума. Только хорошо подготовленные спортсмены способны работать в этой зоне, и то относительно непродолжительные периоды времени. На Рис. 1 представлен пример записи частоты сердечных сокращений с устройства-монитора, сделанной во время тренировочного бега.



Рис. 1. Пример записи с монитора ЧСС во время тренировочного бега

Тренировочная нагрузка характеризуется объемом и интенсивностью. В легкой атлетике интенсивность – это, по сути, скорость бега, и она напрямую соответствует определенным уровням пульсовой активности. Конкретные значения являются в высокой степени индивидуальными, хотя также варьируют в зависимости от физического и психологического состояния и, конечно, уровня тренированности. Так, бегуны-новички показывают около 130 ударов в минуту, хорошо тренированные – до 150 ударов в минуту, а высококлассные спортсмены могут похвастаться пульсом, превышающим 170 ударов в минуту [12]. Очевидна позитивная связь пульсовой активности с уровнем показываемых результатов, хотя, конечно, простое повышение уровня максимального пульса – это неверная стратегия тренировок и развития легкоатлета. Как мы отмечали выше, улучшение способностей спортсмена происходит во время восстановления, и сердце, будучи одним из органов и подчиняясь общим правилам, не является исключением. Таким образом, для лучшего прогнозирования имеет смысл рассматривать не средний достигаемый пульс легкоатлета, а **пульсовые показатели в различных ситуациях**. В частности, такой подход избавляет от необходимости оценивать текущий уровень тренированности спортсмена.

В следующих разделах нашей статьи мы анализируем данные реальных бегунов и проверяем гипотезу, что дисперсия в показываемых ими результатах может быть объяснена рядом факторов, представляющих собой пульсовые показатели при различных видах тренировок.

С. Данные легкоатлетов

В общей сложности мы собрали данные о 14 спортсменах, занимающихся бегом и проживающих в г. Новосибирске (Россия). Среди них было 7 женщин и 7 мужчин, возраст находился в диапазоне от 16 до 35 лет (среднее значение 21,3, дисперсия 2,98). Они добровольно согласились участвовать в исследовании и при помощи разработанной нами гугл-формы предоставили данные о своих пульсовых показателях:

- Пульс при скоростной тренировке (уд/мин): P_{SpT} ;
- Пульс на восстановительном кроссе (уд/мин): P_{RC} ;
- Пульс на рабочем кроссе (уд/мин): P_{WC} ;
- Пульс при силовой тренировке (уд/мин): P_{StT} ;

- Максимальный пульс (уд/мин): P_{max} ;
- Пульс в покое (уд/мин): P_R ;
- Время восстановления пульса (мин): T_{PR} .

Мы также зарегистрировали пол (Gender) и возраст (Age) спортсменов и собрали дополнительную информацию об их текущем тренировочном процессе (объем беговой дистанции в неделю, количество скоростных работ в неделю и т.п.) и питания. Зависимой переменной (Т) являлся результат спортсмена (в секундах) на типовой дистанции в 1 км (мы попросили их пробежать эту дистанцию выкладываясь на 95% от своих способностей).

III. РЕЗУЛЬТАТЫ

A. Описательная статистика

В таб. 1 представлена информация о легкоатлетах, участвовавших в исследовании (средние значения и стандартные отклонения SD приведены в последнем ряду). С целью проверки нормальности распределения зависимой переменной (Т), мы применили тест Колмогорова-Смирнова ($D_{14} = 0,22$; $p = 0,065$) и Шапиро-Уилка ($W_{14} = 0,856$; $p = 0,027$). Поскольку результаты первого теста не позволяют отвергнуть гипотезу о нормальности, мы можем сделать вывод, что отклонение от нормальности является умеренным. Соответственно, в дальнейшем анализе мы будем считать зависимую переменную нормально распределенной, что позволит применять более мощные статистические методы.

ТАБ. 1. ИНФОРМАЦИЯ О ЛЕГКОАТЛЕТАХ

ID	Gender	Age	Training distance per week, km	High-speed runs per week	Running time, s (1 km distance)
C1	M	35	15	2	210
C2	F	16	60	3	195
C3	M	19	27	3	210
C4	F	19	60	2	198
C5	F	20	70	3	205
C6	F	19	50	3	195
C7	F	17	60	3	190
C8	M	24	70	3	158
C9	F	21	20	3	218
C10	F	20	35	3	220
C11	M	24	160	2	150
C12	M	23	200	3	154
C13	M	20	70	2	156
C14	M	21	80	3	155
Среднее (SD)		21,3 (4,58)	69,8 (51,4)	2,71 (0,47)	186,7 (26,3)

ТАБ. 2. ПУЛЬСОВЫЕ ПОКАЗАТЕЛИ ЛЕГКОАТЛЕТОВ

ID	При скоростной тренировке	На восстановительном кроссе	На рабочем кроссе	При силовой тренировке	Максимальный пульс	Пульс в покое	Время восстановления пульса, мин
C1	185	160	165	130	195	60	4
C2	153	164	164	145	200	77	30
C3	150	132	110	96	162	78	3
C4	184	132	150	170	190	52	5
C5	180	150	160	170	190	72	5
C6	185	130	150	150	189	66	7
C7	180	150	165	140	190	48	30
C8	140	100	95	85	160	80	5

ID	При скоростной тренировке	На восстановительном кроссе	На рабочем кроссе	При силовой тренировке	Максимальный пульс	Пульс в покое	Время восстановления пульса, мин
C9	162	108	190	108	108	66	3
C10	170	120	140	90	170	60	3
C11	165	110	140	80	180	48	1
C12	190	120	150	80	200	40	15
C13	188	150	165	115	195	54	2
C14	170	150	160	90	190	55	3
Среднее (SD)	171,6 (15,7)	134,0 (20,4)	150,3 (24,0)	117,8 (32,7)	179,9 (24,4)	61,1 (12,5)	8,29 (9,79)

В таб. 2 представлена описательная статистика по пульсовым показателям (аналогичным образом, средние значения и стандартные отклонения приведены в последнем ряду). Данные позволяют предположить, что пульсовые показатели в разных видах тренировок действительно различаются и, таким образом, потенциально информативны для прогнозирования. Действительно, средний пульс в покое и средний максимальный пульс – т.е. значения, отмечающие границы диапазона пульсовых показателей в целом – различаются почти в 3 раза. Что касается вариабельности по спортсменам, то максимальная была отмечена для переменной «время восстановления пульса» (относительное стандартное отклонение составило 118%).

Далее, мы подсчитали коэффициенты корреляции Пирсона, между T и всеми остальными переменными в нашем исследовании (кроме $Gender$, которая выражена в категориальной шкале). Единственная значимая (при $\alpha = 0,05$) корреляция была найдена для переменной «объем беговой дистанции в неделю» ($r_{12} = -0,756$; $p = 0,002$). Эта закономерность ожидаема, поскольку 1) интенсивная тренировка очевидным образом влияет на результаты спортсмена и 2) более способные спортсмены, как правило, отдают больше усилий тренировкам. Соответствующие корреляции, рассчитанные отдельно для мужчин ($r_5 = -0,726$) и для женщин ($r_5 = -0,748$) оказались весьма схожими. Таким образом, найденная нами значимая закономерность не объясняется наличием двух половых групп в нашем исследовании (хотя, безусловно, тренировочные процессы для мужчин и женщин различаются). Также мы хотели бы отметить, что **по отдельности** ни один из пульсовых показателей не имел статистически значимой корреляции с результатом на дистанции (T).

А. Пульсовые показатели

В рамках более подробного изучения влияния пульсовых показателей, характеристик процесса тренировки и демографических переменных на T , мы применили метод линейной регрессии. Модель, построенная с использованием обратного (backwards)

алгоритма отбора переменных, включила в себя следующие значимые факторы: $Gender$ ($Beta = 0,911$; $p = 0,003$), P_{RC} ($Beta = 0,716$; $p = 0,018$), P_{max} ($Beta = -0,686$; $p = 0,013$), Age ($Beta = 0,471$; $p = 0,038$), and P_{WC} ($Beta = -0,449$; $p = 0,06$). Сама модель также показала высокую статистическую значимость ($F_{5,8} = 7,01$; $p = 0,008$; $R^2 = 0,814$; $R^2_{adj} = 0,698$):

$$T = 189,4 + 46,1Gender + 2,7Age + 0,921P_{RC} - 0,491P_{WC} - 0,738P_{max} \quad (1)$$

При этом, однако, некоторые из промежуточных моделей в ходе применения алгоритма отбора имели более высокий R^2_{adj} . Поэтому мы сравнили информационные критерии Акаике (AIC) для модели с наиболее высоким $R^2_{adj} = 0,760$ и для модели (1):

$$AIC = 2k + n \ln(RSS/n), \quad (2)$$

где k – количество факторов, n – количество наблюдений, а RSS – сумма квадратов остатков в методе наименьших квадратов.

Для модели (1) мы получили $AIC = 76,9$, в то время как для модели-кандидата значение составило $AIC = 73,7$. Соответственно, «потеря информации» в последней ниже, и её следует предпочесть модели (1). Улучшенная модель включила в себя следующие факторы: $Gender$ ($Beta = 1,124$; $p = 0,008$), P_{RC} ($Beta = 1,103$; $p = 0,01$), P_{max} ($Beta = -1,05$; $p = 0,009$), P_{WC} ($Beta = -0,852$; $p = 0,02$), P_{SpT} ($Beta = 0,494$; $p = 0,101$), Age ($Beta = 0,446$; $p = 0,048$), P_{StT} ($Beta = 0,319$; $p = 0,294$):

$$T = 145,6 + 56,9Gender + 2,56Age + 0,826P_{SpT} + 1,42P_{RC} - 0,932P_{WC} - 1,13P_{max} - 0,256P_{StT} \quad (3)$$

Итоговая модель также имела высокую значимость ($F_{7,6} = 6,87$; $p = 0,016$) и несколько более высокий $R^2 = 0,889$ ($R^2_{adj} = 0,760$).

IV. ЗАКЛЮЧЕНИЕ

В нашей статье описано пилотное исследование, направленное на анализ функциональных резервов легкоатлетов и предсказание их способностей. Наши результаты позволяют сделать вывод, что индивидуальные пульсовые показатели, регистрируемые в различных контекстах и при различных типах тренировок, могут выступать предикторами результатов в беге. Теоретическим обоснованием этого является то, что эти показатели, по сути, отражают аэробный и анаэробный метаболизм, обеспечивающий энергию для бега.

Анализируя данные, собранные для 14 легкоатлетов, мы обнаружили единственную статистически значимую корреляцию для беговых результатов – с объемом беговой нагрузки ($r = -0.756$), одной из характеристик процесса тренировки. Хотя эта значимость была ожидаема, этот фактор не имеет особой практической значимости для прогнозирования индивидуальных результатов, т.к. сам является производным от возможностей и резервов спортсмена. А вот в регрессионных моделях (1) и (3), которые мы построили используя обратный алгоритм отбора переменных, важными факторами оказались только пульсовые показатели и демографические характеристики. Интерпретация наиболее влиятельных пульсовых показателей в обеих моделях хорошо согласуется с теорией: более высокий пульс у спортсмена во время восстановительных этапов означает, что его результаты будут ниже, а высокий максимальный пульс обычно означает высокие способности у данного легкоатлета.

Лучшая модель (3), выбранная на основе информационного критерия Акаике, объяснила 88,9% дисперсии в результатах легкоатлетов. Коэффициенты R^2 , полученные в нашем исследовании, имеют довольно высокие абсолютные значения, однако найти базис для сравнения представляется затруднительным, поскольку большинство моделей в предметной области построены на основе подходов машинного обучения к задаче классификации, а не регрессии. В качестве относительно релевантной работы можно упомянуть [10], где также рассматриваются факторы, связанные с сердечным ритмом, и достигается $R^2 = 0.48$, но в рамках задачи избегания травм. Кроме того, в нашем исследовании участвовала гораздо меньшая по объему и, пожалуй, более гомогенная группа спортсменов.

Еще одно ограничение нашего исследования, помимо объема выборки, это относительно ограниченный набор использованных факторов. В частности, мы не рассматривали потенциально релевантные факторы, характеризующие технику бега, психологическую и тактическую подготовленность и т.п. Хотя это было сделано сознательно, исходя из цели нашей статьи, заключающейся в исследовании показателей, связанных с сердечным ритмом, мы можем предположить, что качество моделей могло бы улучшиться от включения дополнительных факторов. Дальнейшие перспективы

наших исследований включают в себя использование пульсовых показателей для разработки рекомендаций по составлению планов тренировок и наиболее подходящих спортсменам дистанций.

Тем не менее, мы полагаем, что уже текущие результаты нашей работы могут представлять интерес для исследователей, специализирующихся в области спортивной медицины, а также для тренеров и для лиц, принимающих решения в управлении спортом. Что же касается легкоатлетов, то мы хотели бы отметить, что улучшение соответствующих пульсовых показателей должно производиться с осторожностью. Планы тренировок должны включать в себя достаточные периоды восстановления, чтобы избежать перетренированности и снизить вероятность травматических последствий.

БЛАГОДАРНОСТИ

Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта 19-29-01017.

БИБЛИОГРАФИЯ

- [1] Куликов И.А. Планирование тренировки в спортивном беге. Сб. труд. межд. конф. Наукоемкие технологии и инновации, Белгород, 2016. С. 198-202.
- [2] Lotfi S. Machine Learning for sport results prediction using algorithms // International Journal of Information Technology and Applied Sciences. 2021. No. 3(3). P. 148-155.
- [3] Prasetyo D. Predicting football match results with logistic regression // IEEE International Conference On Advanced Informatics: Concepts, Theory And Application (ICAICTA), 2016. P. 1-5.
- [4] Horvat T., Job J. The use of machine learning in sport outcome prediction: A review // Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery. 2020. No. 10(5). e1380.
- [5] Rudrapal D. et al. A deep learning approach to predict football match result // Computational Intelligence in Data Mining. Singapore: Springer. 2020. P. 93-99.
- [6] Bakaev M., Avdeenko T. Intelligent information system to support decision-making based on unstructured web data // ICIC Express Letters. 2015. No. 9(4). P. 1017-1023.
- [7] Feely C. et al. Providing explainable race-time predictions and training plan recommendations to marathon runners // Proceedings 14th ACM Conference on Recommender Systems. 2020. P. 539-544.
- [8] Waleriańczyk W., Stolarski M. Personality and sport performance: The role of perfectionism, Big Five traits, and anticipated performance in predicting the results of distance running competitions // Personality and Individual Differences. 2021. No. 169. 109993.
- [9] Liu Q. et al. Classification of runners' performance levels with concurrent prediction of biomechanical parameters using data from inertial measurement units // Journal of Biomechanics. 2020. No. 112. 110072.
- [10] Albert J.A. et al. Using Machine Learning to Predict Perceived Exertion During Resistance Training With Wearable Heart Rate and Movement Sensors // Proceedings IEEE International Conference on Bioinformatics and Biomedicine (BIBM). 2021. P. 801-808.
- [11] Epanchintseva A., Bakaev M. Predicting Performance and Functional Reserves of Athletes Based on Their Pulse Indicators in Different Trainings // Proc. Internet and Modern Society (IMS-2022). St. Petersburg: ITMO University. 2022. (In Print).
- [12] Путилина Т.А., Пасечник Ж.В. Методические особенности подготовки студентов в беге на средние дистанции // Сб. труд. Всерос. конф. Физическое воспитание и спорт: актуальные вопросы теории и практики. 2017. С. 84-87.

Епанчинцева Алина Игоревна, магистрант направления «Биотехнические системы и технологии», Новосибирский государственный технический университет, Россия, (alusik_e@mail.ru).
Бакаев Максим Александрович, кандидат технических наук, доцент, Новосибирский государственный технический университет, Россия, ORCID 0000-0002-1889-0692 (bakaev@corp.nstu.ru).

Modeling runners' performance based on e-monitoring of their heart rate indexes

Alina I. Epanchintseva, Maxim A. Bakaev

Abstract— Today's international sport is a competition of fast managerial decisions, high technology and strong investments. Correspondingly, rational selection of capable sportsmen is crucial for optimal allocation of the limited training resources. In our paper, we perform a pilot experimental study with 14 middle-distance runners and propose a model for predicting performance of athletes that is not based on training process-related factors or on previous performance logs. Instead, we rely on heart rate-related indexes that can be relatively easily monitored using today's e-sensors and mobile devices. In total, we consider 11 factors, but the best model that explains 89% of the variance in performance on the characteristic 1 km distance includes 5 of them (in addition to the demographic factors). Particularly, high pulses at recovery cross and during speed training negatively affect performance, whereas high maximum pulse has significant positive effect. Interestingly, no factors related to the training process, such as the running volume per week, were significant in the model. We believe that our results, even though preliminary, can be of interest to athletes, trainers and sports managers who seek to optimize the training schedules and form balanced national teams.

Keywords— pulse indexes, medical information systems, sports management, running performance.

ACKNOWLEDGMENT

The reported study was funded by RFBR according to the research project No. 19-29-01017.

REFERENCES

- [1] I.A. Kulikov, "Planning of training in sport running", *Proceedings of International Conference on Research-intensive technologies and innovations*, Belgorod, 2016, pp. 198-202. (in Russian)
 - [2] S. Lotfi, "Machine Learning for sport results prediction using algorithms", *International Journal of Information Technology and Applied Sciences*, 3(3), pp. 148-155, 2021.
 - [3] D. Prasetyo, "Predicting football match results with logistic regression", *IEEE International Conference On Advanced Informatics: Concepts, Theory And Application (ICAICTA)*, pp. 1-5, 2016.
 - [4] T. Horvat, J. Job, "The use of machine learning in sport outcome prediction: A review", *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 10(5), e1380, 2020.
 - [5] D. Rudrapal et al., A deep learning approach to predict football match result. In *Computational Intelligence in Data Mining*, Singapore: Springer, 2020, pp. 93-99.
 - [6] M. Bakaev, T. Avdeenko, "Intelligent information system to support decision-making based on unstructured web data", *ICIC Express Letters* 9(4), pp. 1017-1023, 2015.
 - [7] C. Feely et al., "Providing explainable race-time predictions and training plan recommendations to marathon runners", *Proceedings 14th ACM Conference on Recommender Systems*, 2020, pp. 539-544.
 - [8] W. Waleriańczyk, M. Stolarski, "Personality and sport performance: The role of perfectionism, Big Five traits, and anticipated performance in predicting the results of distance running competitions", *Personality and Individual Differences*, vol. 169, 109993, 2021.
 - [9] Q. Liu et al., "Classification of runners' performance levels with concurrent prediction of biomechanical parameters using data from inertial measurement units", *Journal of Biomechanics*, vol. 112, 110072, 2020.
 - [10] J.A. Albert et al., "Using Machine Learning to Predict Perceived Exertion During Resistance Training With Wearable Heart Rate and Movement Sensors", *IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 2021, pp. 801-808.
 - [11] A. Epanchintseva, M. Bakaev, "Predicting Performance and Functional Reserves of Athletes Based on Their Pulse Indicators in Different Trainings", *Proc. Internet and Modern Society (IMS-2022)*, St. Petersburg: ITMO University, in publishing.
 - [12] T.A. Putilina, Z.V. Pasechnik. "Methodological particulars of training students in medium distances running", *Proceedings All-Russian Conference on Physical Education and Sport*, 2017, pp. 84-87. (in Russian)
- Epanchintseva Alina Igorevna**, Master student, major "Biotech Systems and Technologies", Novosibirsk State Technical University, Russia (alusik_e@mail.ru).
Bakaev Maxim Alexandrovich, PhD (Technical Sciences), Associate Professor, Novosibirsk State Technical University, Russia, , ORCID 0000-0002-1889-0692 (bakaev@corp.nstu.ru).