

Salivary Markers for Quantitative Dehydration Estimation During Physical Exercise

Matthias Ring, *Student Member, IEEE*, Clemens Lohmueller, Manfred Rauh, Joachim Mester, and Bjoern M. Eskofier, *Member, IEEE*

Abstract—Salivary markers have been proposed as noninvasive and easy-to-collect indicators of dehydrations during physical exercise. It has been demonstrated that threshold-based classifications can distinguish dehydrated from euhydrated subjects. However, considerable challenges were reported simultaneously, for example, high intersubject variabilities in these markers. Therefore, we propose a machine-learning approach to handle the intersubject variabilities and to advance from binary classifications to quantitative estimations of total body water (TBW) loss. For this purpose, salivary samples and reference values of TBW loss were collected from ten subjects during a 2-h running workout without fluid intake. The salivary samples were analyzed for previously investigated markers (osmolality, proteins) as well as additional unexplored markers (amylase, chloride, cortisol, cortisone, and potassium). Processing all these markers with a Gaussian process approach showed that quantitative TBW loss estimations are possible within an error of 0.34 l, roughly speaking, a glass of water. Furthermore, a data analysis illustrated that the salivary markers grow nonlinearly during progressive dehydration, which is in contrast to previously reported linear observations. This insight could help to develop more accurate physiological models for salivary markers and TBW loss. Such models, in turn, could facilitate even more precise TBW loss estimations in the future.

Index Terms—Dehydration, machine learning, physical exercise, saliva, total body water (TBW).

I. INTRODUCTION

THE normal osmolality¹ of blood plasma is between 280 and 290 mOsm/kg [1, Ch. 4]. Elevations in plasma

osmolality have been shown to track total body water (TBW) loss during physical exercise [2]. The determination of plasma osmolality, however, involves invasive withdrawing of a blood sample and separation of the plasma compartment [3, Ch. 19].

Therefore, salivary osmolality and other salivary markers have been proposed as noninvasive and easy-to-collect alternative for TBW loss tracking [4]–[9]. This proposal was basically motivated by the physiological process that the precursor fluid for salivation is filtrated from plasma [10], and hence, salivary composition should reflect changes in plasma.

Walsh *et al.* first investigated this approach and reported that salivary osmolality, salivary protein concentration, and salivary flow rate are indeed correlated with plasma osmolality [4] and TBW loss [5]. However, they also reported considerable intersubject variability in these markers, which was assumed to cause challenges in the development of subject-independent methods for TBW loss tracking [4], [5].

The potential of salivary osmolality was further investigated in subsequent studies [6]–[9]. Taylor *et al.* [7], for example, also reported considerable intersubject variability in salivary osmolality. Nevertheless, Taylor *et al.* [7] and Muñoz *et al.* [9] showed that threshold-based classifications can distinguish dehydrated from euhydrated subjects. Classifications between more than two hydration conditions, however, did not achieve satisfactory results, which was partly attributed to the intersubject variability [7].

To facilitate TBW loss tracking using salivary markers and to handle the presence of intersubject variability, we investigated a machine-learning approach based on relative changes in salivary markers. We also included additional salivary markers to minimize remaining effects of intersubject variability. These markers were amylase, chloride, cortisol, cortisone, and potassium. Most of them have been examined regarding various aspects of physical exercise (e.g., [11]) but not yet regarding TBW loss tracking [4]–[9]. Therefore, we first applied an exhaustive feature selection to identify those markers that best track TBW loss. The identified markers were then presented to regression methods that learned to estimate TBW loss quantitatively, as an advancement to previous binary classifications [4]–[9]. This approach, when implemented on wearable devices that can measure salivary markers in the field (e.g., [12]–[14]), could facilitate wearable systems that continuously and noninvasively track TBW loss in the field.

Finally, our analysis also indicated that salivary markers increase nonlinearly during progressive dehydration. This insight

Manuscript received February 29, 2016; revised June 6, 2016 and August 4, 2016; accepted August 5, 2016. Date of publication August 10, 2016; date of current version September 1, 2017. The work of M. Ring was supported by a research fellowship from Deutsche Telekom Stiftung.

M. Ring and B. M. Eskofier are with the Digital Sports Group, Pattern Recognition Laboratory, Friedrich-Alexander-Universität Erlangen-Nürnberg, 91058 Erlangen, Germany (e-mail: matthias.ring@cs.fau.de; bjoern.eskofier@fau.de).

C. Lohmueller is with the Department of Surgery, Friedrich-Alexander-Universität Erlangen-Nürnberg, 91058 Erlangen, Germany (e-mail: clemens.lohmueller@uk-erlangen.de).

M. Rauh is with the Department of Paediatrics, Friedrich-Alexander-Universität Erlangen-Nürnberg, 91058 Erlangen, Germany (e-mail: manfred.rauh@uk-erlangen.de).

J. Mester is with the Institute of Training Science and Sport Informatics, German Sport University Cologne, 50933 Cologne, Germany (e-mail: Mester@dshs-koeln.de).

Digital Object Identifier 10.1109/JBHI.2016.2598854

¹Osmolality denotes the number of moles of osmotically active solutes per weight of solution. Osmotically active solutes in plasma are, for example, sodium, chloride, and proteins.

might help to understand some challenges that previous linear approaches [4], [5], [7], [9] encountered. It might also facilitate the development of more accurate physiological models for salivary markers and TBW loss in the future.

II. METHODS

This section first describes the study that was conducted to collect salivary samples as well as reference values of TBW loss during physical exercise. Then, the machine-learning approach for TBW loss tracking and its evaluation on the collected dataset are described.

A. Data Collection Study

The data collection study included various other physiological markers besides salivary markers. The following description, however, recalls only details that are relevant for this paper.²

1) Subjects: Ten male subjects (S1, ..., S10) volunteered to participate in the study. Mean and standard deviation of height, body weight, and age were 179 ± 7.5 cm, 79.3 ± 9.0 kg, and 25.5 ± 3.7 years, respectively. All subjects provided written informed consent after the study protocol was approved by the local ethics committee. Individuals with considerably different age, body composition, or diseases that might affect salivary composition were not considered in this study to minimize confounding variables and side effects.

2) Preliminary Examination: At least one day and at most one week before the data collection, subjects underwent a preliminary examination.³ This examination included determining the ventilatory threshold and (if possible) maximum oxygen uptake of every subject. The values were calculated according to Scharhag-Rosenberger [15] based on the subjects' performance at an incremental exercise test to volitional exhaustion on a treadmill.

3) Experimental Procedures: Four preconditions were defined to ensure comparable hydration conditions among the individual subjects. First, subjects were asked to refrain from strenuous physical activity, alcohol, and caffeine on the day before the data collection. Second, subjects were asked to report to the laboratory at 6:30 A.M. on the day of the data collection, following a 10-h overnight fast. Third, subjects received an identical breakfast in the laboratory: 250 ml of apple juice mixed with water after arrival; 312 ml of a meal replacement drink (Fit and Feelgood Diät-Shake, Layenberger, Rodenbach, Germany) 1 h later. Fourth, subjects were not allowed to consume any foods or beverages until the end of the data collection.

The collection of salivary markers as well as reference values of TBW loss started at 9:35 A.M. For this purpose, subjects ran for a total of 120 min. To minimize confounding environmental effects, running was performed on a treadmill that was placed in a laboratory, instead of running outdoors. To minimize confounding effects of clothing, subjects wore identical

t-shirts and shorts (Response 3-Stripes, Adidas, Herzogenaurach, Germany). To minimize confounding effects of individual physical capacity, subjects ran at individual speeds that corresponded to their min {ventilatory threshold, 60% maximum oxygen uptake}.

For the collection of measurements, the 120 min of running were partitioned into eight intervals. Every interval consisted of 15 min of running and 8 min of resting. Measurements were collected during the 8 min of resting. In addition, baseline measurements were collected immediately before the first running interval began. In the following, measurements will be referred to using an index i , with $i = 0, \dots, 8$, where $i = 0$ denotes baseline measurements, $i = 1$ denotes measurements after the first 15-min running interval, etc.

4) Salivary Markers: Saliva was obtained using Salivette cotton tubes (Sarstedt, Nürnbrecht, Germany), which were positioned under the tongue during resting. These collectors were also used in three [4], [5], [7] of the six previous studies on saliva and TBW loss. Mouth rinses before salivary collection were not used because it was observed that mouth rinses alter salivary osmolality for about 15 min [16]. In the laboratory, salivary samples were then analyzed for seven markers: amylase, chloride, cortisol, cortisone, osmolality, potassium, and proteins.

Amylase is an enzyme to breakdown starch and is activated by chloride ions [3, Ch. 24]. Cortisol and cortisone are steroid hormones, which affect, for example, blood glucose concentration [17, Ch. 77]. Potassium is found in saliva as a result of the salivation process [17, Ch. 64]. Proteins denotes the total concentration of all proteins found in saliva.

Amylase concentration (α -amylase) was measured with the enzymatic colorimetric method using the substrate 4,6-ethylidene-*p*-nitrophenyl- α -D-maltoheptaoside from Roche Diagnostics (Mannheim, Germany). The diluted salivary samples (1 + 999) were analyzed with a Roche integra system 800. Chloride concentration was analyzed with a chloridometer (CM20, Kreienbaum, Langenfeld, Germany). Cortisone and cortisol concentrations were determined by liquid chromatography tandem mass spectrometry. Measurements of potassium (ISE, using automatically diluted specimens) and proteins (turbidimetric, benzethoniumchlorid) were performed on a Cobas Integra 800 (Roche Diagnostics, Mannheim, Germany). Osmolality was analyzed with a vapor pressure osmometer (Vapro, Wescor, Logan, UT, USA).

All salivary markers were determined using laboratory equipment, though practical applications would require wearable devices that can measure the markers in the field. However, first prototypes of such devices have already been reported in recent literature, e.g., [12]–[14], and should be employed in follow-up studies.

5) TBW Loss: Reference values of TBW loss were obtained by setting TBW loss after every running interval equal to body weight loss after every running interval. This method is based on the assumption that if there is no food or beverage intake, and no urine and fecal losses, then body weight during physical exercise only changes because of water loss [18]. This method was also used in all previous studies on saliva and TBW

²The full study protocol may be found in the International Clinical Trials Registry Platform of the World Health Organization using the identifier DRKS00005301.

³Details may be found in the full study protocol (see footnote 2).

TABLE I
OVERVIEW OF MISSING VALUES IN SALIVARY MARKERS

	Missing values ^a per subject ^b				
	S1	S2	S3, ..., S8	S9	S10
Amylase	0, 1, 3	2, 5, 6, 7, 8	–	0	–
Chloride	–	4, 6	–	0	–
Cortisol	0, 1, 7	–	–	0	–
Cortisone	0, 1, 7	–	–	0	–
Osmolality	–	6	–	0	–
Potassium	0, 1	2, 5, 6, 7, 8	–	0	–
Proteins	–	5, 6, 7, 8	–	0	–

^a Missing values are indicated using the index i , with $i = 0, \dots, 8$, which denotes the number of the 15-min running interval after which the corresponding salivary sample was collected. That is, $i = 0$ denotes baseline samples, $i = 1$ denotes samples after the first 15-min running interval, etc.

^b Ten subjects, denoted with S1, ..., S10, participated in the study.

loss [4]–[9]. Therefore, subjects were asked to undress and to remove all sweat on their body using a towel. Then, nude body weight was measured with a high precision weight scale (± 5 g accuracy; DE 150K2D, Kern & Sohn, Balingen-Frommern, Germany). The difference in nude body weight (in kg) between two consecutive running intervals was then set equal to TBW loss (in l) after the corresponding running interval. Urine and fecal losses did not occur during the running procedure and were, therefore, not considered as further loss of TBW.

B. TBW Loss Estimation Using Salivary Markers

1) Dataset: The measured values for all seven salivary markers as well as reference values of TBW loss were compiled into one dataset. This yielded a dataset with some missing values, because isolated salivary samples contained insufficient amounts of saliva for measuring one, or multiple, markers in the laboratory (see Table I).

For a tradeoff between excluding subjects with missing values and having still enough subjects for an informative evaluation of the machine-learning algorithms (leave-one-subject-out cross-validation [LOSO-CV]; see below), we decided to exclude S1 and S2 but to include S9. For this purpose, S9's missing values at $i = 0$ were extrapolated using a linear regression (LR) of the following three values at $i = 1, 2, 3$. The linear extrapolation was selected based on the visual observation that the markers increased rather constantly during the first three running intervals compared to the subsequent running intervals, if averaged over all subjects (see also parallel coordinate plot below).

2) Preprocessing: Previous studies [4], [5], [7] reported considerable intersubject variability in salivary markers. Therefore, we used relative changes instead of absolute values. Relative changes can also be motivated by the inter- and intrasubject variability in salivary osmolality that was observed over several consecutive days [19]. Relative changes were supposed to minimize effects of intersubject variability in subsequent processing steps. For this purpose, all seven salivary markers were converted into relative changes with respect to their baseline values at $i = 0$.

TABLE II
CORRELATION BETWEEN RELATIVE CHANGES IN SALIVARY MARKERS AND ABSOLUTE TBW LOSS

	Pearson	Spearman
Amylase	0.60 ($p < 0.001$)	0.62 ($p < 0.001$)
Chloride	0.84 ($p < 0.001$)	0.91 ($p < 0.001$)
Cortisol	0.35 ($p = 0.002$)	0.42 ($p < 0.001$)
Cortisone	0.59 ($p < 0.001$)	0.66 ($p < 0.001$)
Osmolality	0.59 ($p < 0.001$)	0.77 ($p < 0.001$)
Potassium	0.68 ($p < 0.001$)	0.86 ($p < 0.001$)
Proteins	0.65 ($p < 0.001$)	0.69 ($p < 0.001$)

Reference values of TBW loss were not converted into relative changes. The motivation for this decision was that the primary application of the present approach would be the estimation of TBW loss during physical exercise. This would enable recommendations on how much fluid should be consumed. The estimation of relative TBW loss, however, would be inappropriate in such situations. Relative values could not be converted into concrete volumes of fluid that should be consumed, because the absolute volume of baseline TBW is typically unknown in field applications.

3) Regression Method: A preliminary data analysis was conducted to explore which regression method might best estimate TBW loss. For this purpose, Pearson's [20] as well as Spearman's [21] correlation coefficients were computed between every relative salivary marker and TBW loss. Additionally, relative salivary markers as well as TBW loss were averaged over all subjects and depicted in a parallel coordinate plot [22]. Finally, osmolality and proteins were plotted for every subject because these two measurements were mainly investigated in previous studies [4]–[9].

Spearman's coefficient, which indicates generic monotonic correlation, showed larger values than Pearson's coefficient, which indicates linear correlation (see Table II). This observation indicated that correlations are present, but these correlations might not necessarily be of linear nature. The parallel coordinate plot (see Fig. 1) as well as the osmolality and proteins plots (see Figs. 2 and 3, respectively) supported this observation visually. The trend over time of the salivary markers exhibited nonlinear, but also monotonic, characteristics. In contrast, the trend over time of TBW loss exhibited linear characteristics.

As a result, the preliminary data analysis suggested nonlinear regression methods for TBW loss estimation, in contrast to the previously applied linear approaches [4]–[9]. We selected a Gaussian process regression (GPR) for this purpose (see [23]). GPRs are able to learn nonlinear characteristics, which seem to be present in our dataset. GPRs are flexible and data-driven, they do not require a fully prespecified model, like a quadratic function where only the coefficients are free modeling parameters. Recently, GPRs have also been shown to model and process various physiological data successfully (e.g., [24]–[27]).

GPR was configured identically in all of the following experiments. It was configured to use a linear mean function, a squared exponential covariance function, and a Gaussian likelihood function. The parameters of both the mean function and covariance function were initialized with 1. The parameter of the

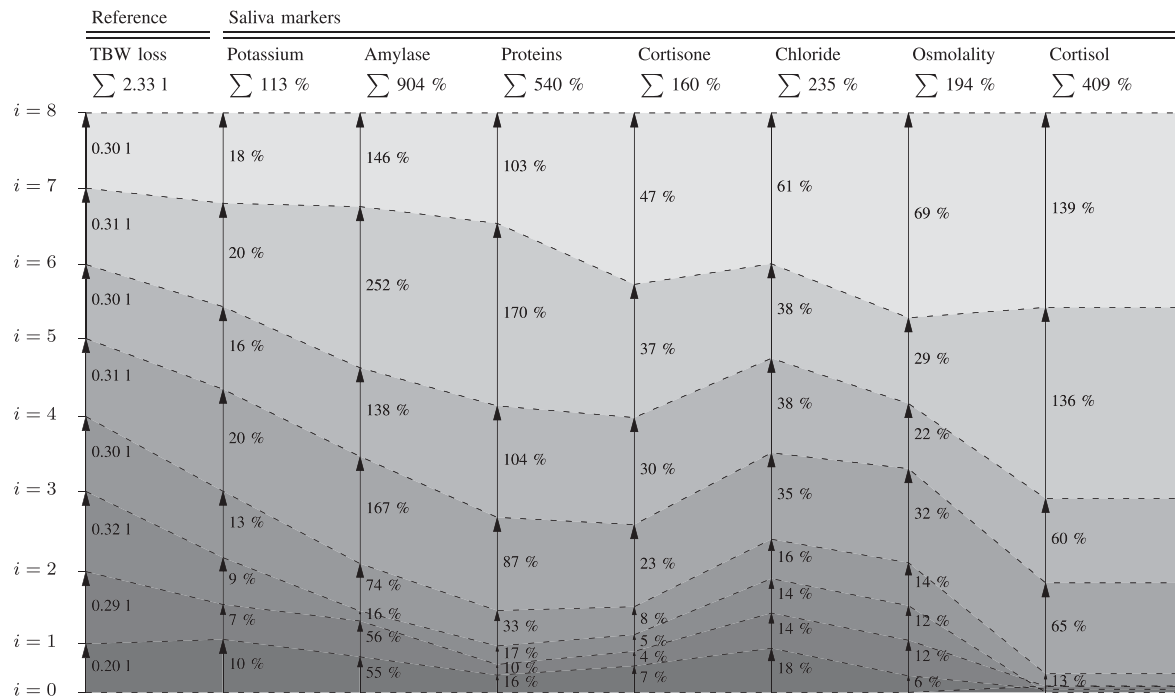


Fig. 1. Parallel coordinate plot depicting absolute TBW loss as well as relative salivary markers (left to right) and their progression over time (bottom to top). Values are means over all subjects and relative to baseline. For example, absolute TBW loss was 0.2 L after the first running interval, 0.49 L (0.20 L + 0.29 L) after the first two running intervals, until a total TBW loss of 2.33 L was reached at the end of the running workout. Similarly, potassium concentration was 10% increased compared to baseline after the first running interval, 17% (10% + 7%) after the first two running intervals, until a total increase of 113% compared to baseline was reached at the end of the running workout. Note that the first three values for cortisol were omitted because of clarity of presentation. These values were -2% , 2% , and -4% . The plot illustrates that TBW loss increased almost linearly during progressive dehydration, with about 0.3 L per interval. In contrast, salivary markers increased rather nonlinearly during progressive dehydration. They grew slowly in the beginning ($i = 1$ to $i = 4$) but increased rather dramatically as dehydration progressed ($i \geq 4$). No salivary marker seems to be related to TBW loss by a straightforward linear correlation. This is illustrated by the horizontal lines, which connect all values that were collected at the same time. These lines illustrate the nonlinearity because linearly correlated values would appear on common horizontal levels in this plot, i.e., the horizontal lines would be straight if the salivary markers would be linearly correlated to TBW loss and among each other.

likelihood function was initialized with 0.1. Then, the parameters were optimized by minimizing the negative log marginal likelihood with a conjugate gradient descent (maximum of 1000 iterations).

For comparisons between GPR and previous, linear approaches [4]–[9], LR was also included using the least-squares method for training [28, Ch. 3].

Besides GPR, further nonlinear regression methods would have been, for example, support vector regression (SVR) [29], [30] or random forest regression [31]. The reason for disregarding such methods was motivated by the training data. SVR parameters are usually determined by combining a grid search with cross-validation, whereas random forests often rely on bootstrapping to build the individual trees. These approaches, therefore, would have required further partitioning or resampling of the training data, within an already nested cross-validation that was required to evaluate the feature selection (see below). This, in turn, would have decreased the amount of training data and probably caused more instabilities. Therefore, we selected GPR, whose gradient descent for parameter optimization does not require further partitioning or resampling of the training data.

4) Machine Learning: Both regression methods, GPR and LR, were evaluated in combination with a preceding feature selection. Feature selection was included in case not all mea-

sured salivary markers are essential for accurate TBW loss estimation.

The feature selection was implemented using a k -exhaustive search [32, Ch. 7.1]. The exhaustive search examined all possible combinations of selecting k out of all seven markers. The parameter k was increased from $k = 1$, i.e., finding the best single salivary marker, up to $k = 7$, i.e., using all salivary markers.

The performance of every combination of salivary markers was assessed using a LOSO-CV [28, Ch. 7]. The LOSO-CV method was chosen because it reasonably simulates future application scenarios [33], in which the present approach would be employed for estimating TBW loss of, most probably, unknown subjects.

The LOSO-CV marked every subject once as held-out subject, whereas the remaining subjects were marked as training subjects. Three steps were then performed for every held-out subject. First, the regression was trained using the data of the training subjects after every 15-min running interval. These data included only the current combination of k salivary markers as well as reference values of TBW loss. Second, TBW loss was estimated for the held-out subject after every 15-min running interval. This estimation was also performed using only the current combination of k salivary markers. Third, the absolute error between estimated TBW loss and reference TBW loss was averaged over all 15-min running intervals.

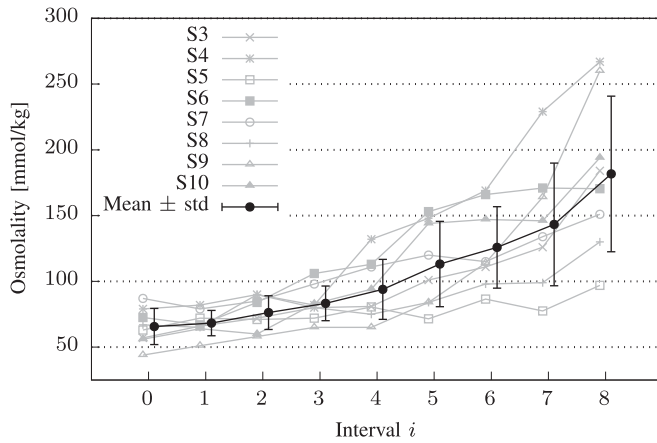


Fig. 2. Osmolality after every 15-min running interval, $i = 1, \dots, 8$, as well as baseline osmolality before physical exercise, $i = 0$. Subjects' individual values are shown in gray color; mean and standard deviation over all subjects are shown in black color. For better visualization, individual values have been slightly shifted to the left, whereas mean and standard deviation have been slightly shifted to the right.

The LOSO-CV averaged the individual mean errors over all held-out subjects. This final mean error was returned to the feature selection as performance measure for the current combination of k salivary markers.

For every value of k , the feature selection eventually selected the combination of k salivary markers that achieved the smallest error. The final regressions were then trained using these selected salivary markers and the data of all subjects that were presented to the LOSO-CV.

The performance of the final regressions was also evaluated using a LOSO-CV. This outer LOSO-CV received all subjects, in contrast to the above described inner LOSO-CV that ultimately received all but the held-out subject of the outer LOSO-CV.

The outer LOSO-CV obtained the final regression for every held-out subject by running the feature selection and inner LOSO-CV on its training subjects, as described above. Then, the outer LOSO-CV tested the final regression on its held-out subject. The results of these TBW loss estimations and the selected features for every held-out subject were recorded. These results are reported below (see Section III) and form the basis for discussions about our entire approach (see Section IV).

III. RESULTS

Fig. 4 depicts the absolute error for the different configurations of the machine-learning approach. The overall smallest error of 0.34 ± 0.10 L was achieved if GPR was employed and five salivary markers were selected in the feature selection. LR achieved its smallest error of 0.36 ± 0.13 L if the best single salivary marker ($k = 1$) was selected in the feature selection.

For the configuration that achieved the smallest error (GPR, $k = 5$), Fig. 5 depicts the progression over time of estimated TBW loss and reference TBW loss. The discrepancy between the means of both quantities was minimal, which indicated that this configuration seems to provide an unbiased estimation of TBW loss.

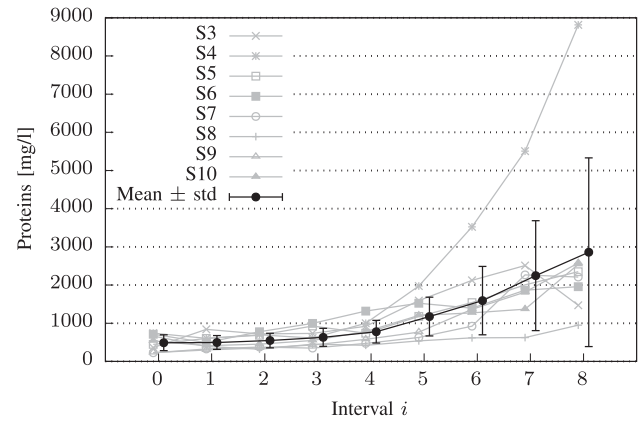


Fig. 3. Same as Fig. 2 for proteins.

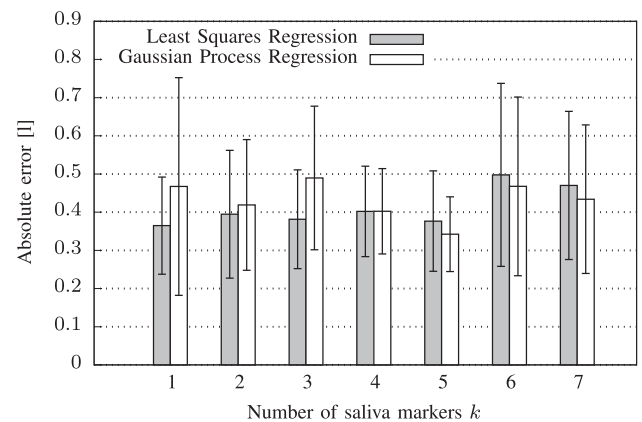


Fig. 4. Absolute error of TBW loss estimation, illustrated per number of salivary markers k that were used for estimation. The figure depicts mean and standard deviation after averaging over all held-out subjects in the outer LOSO-CV.

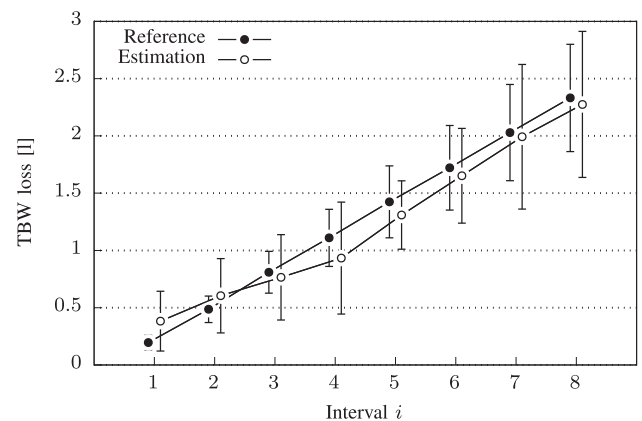


Fig. 5. TBW loss over all 15-min running intervals if GPR was employed and $k = 5$ salivary markers were selected in the k -exhaustive feature selection. The figure depicts mean and standard deviation after averaging over all held-out subjects at every 15-min running interval. For better visualization of the error bars, white-filled points have been slightly shifted to the right, and black-filled points have been slightly shifted to the left.

TABLE III
SALIVARY MARKERS SELECTED IN THE k -EXHAUSTIVE FEATURE SELECTION

k	Regression ^a	Salivary markers ^b
1	LR	chloride (8)
	GPR	chloride (7), proteins (1)
2	LR	chloride (8), osmolality (7), amylase (1)
	GPR	chloride (8), osmolality (3), potassium (2), proteins (2), amylase (1)
3	LR	chloride (8), cortisone (6), osmolality (6), cortisol (2), potassium (2)
	GPR	chloride (8), osmolality (4), potassium (3), proteins (3), amylase (2), cortisol (2), cortisone (2)
4	LR	chloride (8), osmolality (8), cortisol (5), cortisone (4), proteins (3), amylase (2), potassium (2)
	GPR	chloride (8), osmolality (8), cortisone (4), potassium (4), cortisol (3), proteins (3), amylase (2)
5	LR	chloride (8), osmolality (8), cortisol (6), cortisone (6), proteins (6), amylase (4), potassium (2)
	GPR	chloride (8), osmolality (8), cortisol (6), cortisone (6), potassium (5), proteins (4), amylase (3)
6	LR	chloride (8), cortisone (8), osmolality (8), potassium (7), cortisol (6), proteins (6), amylase (5)
	GPR	chloride (8), cortisone (8), cortisol (7), osmolality (7), proteins (7), potassium (6), amylase (5)
Σ	LR	chloride (48), osmolality (37), cortisone (24), cortisol (19), proteins (15), potassium (13), amylase (12)
	GPR	chloride (47), osmolality (30), cortisone (20), potassium (20), proteins (20), cortisol (18), amylase (13)

^a Regression method: linear regression (LR), Gaussian process regression (GPR).

^b Name of salivary marker followed by the number of LOSO-CV folds, in which the marker was selected. For example, if $k = 1$ and LR was employed, chloride was selected in all eight LOSO-CV folds. If $k = 1$ and GPR was employed, chloride was selected in seven of eight LOSO-CV folds and proteins were selected in one of eight LOSO-CV folds.

Table III depicts the salivary markers that were selected in the different feature selections of the outer LOSO-CV. For the configuration that achieved the smallest error (GPR and $k = 5$), the five mostly selected markers were chloride, osmolality, cortisol, cortisone, and potassium. Interestingly, this set was almost identical with two further marker sets. First, the five markers that were mostly selected in all k -exhaustive feature selections (see Table III; Σk , GPR) were chloride, osmolality, cortisone, potassium, and proteins (instead of cortisol). Second, the five markers that showed the highest correlation according to Spearman's coefficient (see Table II) were also chloride, osmolality, cortisone, potassium, and proteins (instead of cortisol).

LR achieved its smallest error if the feature selection searched for the best single marker. This marker was chloride in every LOSO-CV fold. Interestingly, chloride also exhibited the highest linear correlation with TBW loss, according to Pearson's coefficient (see Table II).

IV. DISCUSSION

We explored a machine-learning approach for TBW loss tracking based on salivary markers. In situations like our data collection study, e.g., no fluid intake, severe dehydrations can appear. For such situations, previously investigated markers as well as additional markers were examined for their capability to facilitate quantitative TBW loss estimations. Experiments showed that unbiased estimations, within an absolute error of 0.34 L, are possible by processing salivary markers with non-linear machine-learning methods.

A. Putting the Error into Context

To the best of our knowledge, there is only one other method that noninvasively estimated TBW loss during physical exercise [34]. This method combined bioimpedance with temperature measurements and achieved an error of 0.46 L. Compared to this result, the present error of 0.34 L provides an improvement of 26%. The so-called water-deficit equation [35], which is an invasive method based on plasma analysis, was recently shown to underestimate TBW loss after physical exercise by more than 1.5 L [36]. The water-deficit equation, however, utilized other processing methods and evaluated less markers, i.e., less information. These two differences could explain the superiority of the present saliva-based approach, although saliva is physiologically only a filtrate of blood and saliva analysis is conceptually only a surrogate for blood analysis.

In the broader context, if a subject consumes the amount of fluid estimated by our approach, the possible error of 0.34 L can lead to two situations. First, the subject might consume 0.34 L of fluid more than necessary, assuming that the subject was euhydrated before physical exercise. This situation should not be critical, since moderate hyperhydrations are generally less associated with health problems [37]. Second, the subject might consume 0.34 L of fluid less than necessary, assuming that the subject was euhydrated before physical exercise. This situation should neither be critical, because an increase of 5 mmol/kg in plasma osmolality is usually stated as the threshold where the body activates mechanisms for water retention and acquisition [38]. An increase of 5 mmol/kg in plasma osmolality, however, corresponds to a TBW loss of about 1.4 L, in case of

an average 70-kg body [38]. Thus, the subject, although having a possible deficit of 0.34 L, would have a TBW loss clearly below 1.4 L. In other words, TBW loss estimations by our approach should bring the subject into a hydration state that is tolerated by the body and far from thresholds where mechanisms for water retention and acquisition are activated.

B. Best Set of Salivary Markers

We chose a LOSO-CV for the evaluation because it provides a reasonable estimation of the expected error on future, unknown subjects [33]. Based on the present data, this expected error will be minimal if GPR is employed as regression method and five salivary markers are selected with an exhaustive feature selection.

The LOSO-CV method, however, did not simultaneously determine a final set of five markers. Instead, it determined one set of five markers in every fold, and these sets were not identical over all folds (see Table III; $k = 5$, GPR).

There are two common approaches to decide for a final set in such situations. First, running an additional five-exhaustive feature selection on all subjects, i.e., omitting the outer LOSO-CV. Second, evaluating the different sets from all folds, for example, by choosing the five markers that were mostly selected over all folds.

For the present dataset, we recommend the second approach. There was no stable set of five markers that was clearly selected in the majority of all folds (see Table III; $k = 5$, GPR). This fact prefigured that an additional five-exhaustive feature selection on all subjects might similarly determine a possibly unstable marker set. Therefore, we suggest to define the final set as the five, mostly selected markers over all folds. These markers were chloride, osmolality, cortisol, cortisone, and potassium (see Table III; $k = 5$, GPR). This marker set should represent a more stable selection. The fact that it was almost identical with two other interesting marker sets (see Section III) supports this choice.

Nevertheless, future studies should be conducted to collect more training data and to achieve more stable feature selections. The present work laid the foundation for further research into this direction by demonstrating that TBW loss estimations are possible within an error of, roughly speaking, a glass of water.

C. Advantages of the Machine-Learning Approach

The minimum error was obtained with a nontrivial machine-learning approach (GPR combined with an exhaustive feature selection). This result suggested that a machine-learning approach will also be a reasonable choice in future large-scale studies for selecting the best markers as well as training the regression. The necessity of a feature selection was also emphasized by the result that the error first decreased with respect to the number of markers (see Fig. 4; GPR, $k = 1$ to 5) and then increased again (see Fig. 4; GPR, $k = 5$ to 7). The minimum error was neither achieved by selecting the best single marker nor by employing all available markers simultaneously.

The increasing error for $k > 5$ (see Fig. 4; GPR and LR), however, could also have emerged because of the comparatively

high number of salivary markers compared to the number of subjects. Therefore, future large-scale studies, which will provide a better ratio between the number of salivary markers and the number of subjects, might find that more than five markers are best for TBW loss estimations. The proposed machine-learning approach could provide the fundamental tool to investigate such issues in future studies.

The best LR configuration was outperformed by the best GPR configuration (see Fig. 4 and Section III). Nevertheless, the error of the best LR configuration might still seem acceptable, and the fact that only one marker (chloride) was required might seem attractive in the context of practical implementations on low-resource wearable devices. However, there are two further aspects that should be considered in this context. The first aspect is the analysis of the dataset. The correlation analysis for chloride (see Table II) showed a higher value in Spearman's than in Pearson's coefficient. As argued above (see Section II-B3), this fact indicated the presence of nonlinear correlations, which might not be completely conceivable by an LR. The parallel coordinate plot (see Fig. 1) confirmed this argument visually. The chloride concentration increased rather nonlinearly: 16–18% in the first four 15-min intervals, 35–38% in the following three 15-min intervals, and 61% in the final 15-min interval.

The second aspect is the intersubject variability, which was reported in previous studies [4], [5], [7]. GPR might provide another crucial advantage over LR in the light of this variability. GPR stores the training data and employs them in the prediction process [23, Ch. 2]. This technique gives GPR the opportunity to learn and recall intersubject differences. This could be a crucial advantage in the present application. LR, in contrast, discards the training data once a set of parameters (slope and offset), which provides the best tradeoff to describe all training data at once, has been determined [28, Ch. 3]. Furthermore, there are also optimized GPR variants that can handle large training datasets [23, Ch. 8], in case the present approach will be applied in future large-scale studies.

On basis of the current dataset, its analysis, and the arguments above, we would favor GPR if one approach has to be selected. But we are also aware that future studies with more subjects might reveal other aspects that will have to be considered.

D. Nonlinear Characteristics of Salivary Markers

The present data suggested that the salivary markers increase nonlinearly during progressive dehydration (see Section II-B3). This suggestion is in contrast to the study from Walsh *et al.* [5]. They observed linear correlations of 0.94 between osmolality and relative TBW loss, and 0.97 between proteins and relative TBW loss.

The discrepancy could be explained by two methodological differences. First, Walsh *et al.* [5] calculated correlations between absolute values of salivary markers and relative values of TBW loss. In contrast, we worked with relative values of salivary markers and absolute values of TBW loss, as motivated above (see Section II-B2). Second, and more important, Walsh *et al.* [5] collected only three salivary samples during physical exercise, whereas we collected eight salivary samples. The

present data, therefore, should give a more comprehensive perspective on the progression of salivary markers during physical exercise.

Regarding correlation coefficients, comparisons to the other studies [4], [6]–[9] are not meaningful or not possible. Walsh *et al.* [4] and Muñoz *et al.* [9] did not report correlations between salivary markers and TBW loss, Smith *et al.* [6] computed correlations only between pre-exercise and post-exercise values, and Taylor *et al.* [7] and Horn *et al.* [8] allowed fluid consumption during physical exercise, which might have caused confounding effects.

The nonlinear characteristics were particularly apparent at the transition from interval $i = 4$ to $i = 5$ (see Fig. 1). The percentage growth per interval increased dramatically between these two intervals. For example, chloride increased 14–18% in each of the first four intervals, but in the fifth interval, the growth was doubled to 35%. Cortisone increased 4–8% in each of the first four intervals, but in the fifth interval, the growth was tripled to 23%. Similar changes can be observed in the other markers. Future research could investigate why the changes occurred at this time point, after a TBW loss of about 1 L ($0.20\text{ L} + 0.29\text{ L} + 0.32\text{ L} + 0.30\text{ L}$; see Fig. 1) was reached.

E. Future Research Directions

The discussion thus far suggested that further research would be profitable. As detailed above, future research could stabilize feature selection by collecting more data (see Section IV-B), reveal a clearer picture whether linear or nonlinear regressions perform better (see Section IV-C), or investigate the reasons for the dramatic change in the salivary markers after about 1 L of TBW loss (see Section IV-D). Future research should further investigate physical exercise longer than 120 min, physical exercise with water consumption, and subjects with different age, body composition, or diseases that might affect salivary composition.

Salivary markers could have been influenced by side effects like drying mucous membranes. If future studies observe such effects, appropriate pre-processing algorithms (e.g., principal component analysis [28, Ch. 14.5]) could be able to remove, or minimize, them. Irrespective of whether such effects are present or not, and irrespective of the physiological mechanisms that led to the salivary composition that has been observed in this study, the present approach can already be used in the described setting, because the evaluation showed that accurate TBW loss estimations are possible with salivary markers that have been collected in this very setting.

V. CONCLUSION

This paper explored a machine-learning approach for TBW loss tracking during physical exercise. Previously investigated salivary markers as well as additional, unexplored salivary markers were employed for this purpose. The evaluation demonstrated that quantitative TBW loss estimations are possible within an error of 0.34 L, roughly speaking, a glass of water. This is an advancement to previous binary classifications between dehydrated and euhydrated subjects based on salivary mark-

ers. The results also suggested further investigation of advanced computational methods like the machine-learning approach, because of the intersubject variability and the nonlinearity of the salivary markers. Next steps could include experiments with fluid replacement, which would probably lead to less severe, threatening dehydrations, but in turn, reduce the number of controlled variables (water intake). In the future, such methods, when implemented on wearable devices that can measure salivary markers in the field, could facilitate wearable systems for continuous and noninvasive TBW loss tracking.

ACKNOWLEDGMENT

The authors would like to thank adidas AG and iQ-Move GmbH for support during data collection, and the anonymous reviewers for their valuable comments.

REFERENCES

- [1] Panel on Dietary Reference Intakes for Electrolytes and Water, *Dietary Reference Intakes for Water, Potassium, Sodium, Chloride, and Sulfate*. Washington, DC, USA: Nat. Acad. Press, 2005.
- [2] L. A. Popowski, R. A. Oppliger, G. P. Lambert, R. F. Johnson, A. K. Johnson, and C. V. Gisolfi, "Blood and urinary measures of hydration status during progressive acute dehydration," *Med. Sci. Sports Exercise*, vol. 33, no. 5, pp. 747–753, 2001.
- [3] G. J. Tortora and B. H. Derrickson, *Principles of Anatomy and Physiology*, 12th ed. Hoboken, NJ, USA: Wiley, 2009.
- [4] N. P. Walsh, S. J. Laing, S. J. Oliver, J. C. Montague, R. Walters, and J. L. J. Bilzon, "Saliva parameters as potential indices of hydration status during acute dehydration," *Med. Sci. Sports Exercise*, vol. 36, no. 9, pp. 1535–1542, 2004.
- [5] N. P. Walsh, J. C. Montague, N. Callow, and A. V. Rowlands, "Saliva flow rate, total protein concentration and osmolality as potential markers of whole body hydration status during progressive acute dehydration in humans," *Archives Oral Biol.*, vol. 49, no. 2, pp. 149–154, 2004.
- [6] D. Smith, I. Shalmyeva, J. DeBlois, and M. Winke, "Use of salivary osmolality to assess dehydration," *Prehospital Emergency Care*, vol. 16, no. 1, pp. 128–135, 2011.
- [7] N. A. S. Taylor *et al.*, "Observations on saliva osmolality during progressive dehydration and partial rehydration," *Eur. J. Appl. Physiol.*, vol. 112, no. 9, pp. 3227–3237, 2012.
- [8] G. P. Horn, J. DeBlois, I. Shalmyeva, and D. L. Smith, "Quantifying dehydration in the fire service using field methods and novel devices," *Prehospital Emergency Care*, vol. 16, no. 3, pp. 347–355, 2012.
- [9] C. Muñoz *et al.*, "Assessment of hydration biomarkers including salivary osmolality during passive and active dehydration," *Eur. J. Clin. Nutrition*, vol. 67, no. 12, pp. 1257–1263, 2013.
- [10] J. Thaysen, N. Thorn, and I. Schwartz, "Excretion of sodium, potassium, chloride and carbon dioxide in human parotid saliva," *Amer. J. Physiol.*, vol. 178, no. 1, pp. 155–159, 1954.
- [11] J. Chicharro, A. Lucía, M. Pérez, A. Vaquero, and R. Ureña, "Saliva composition and exercise," *Sports Med.*, vol. 26, no. 1, pp. 17–27, 1998.
- [12] R. Gunter, W. Delinger, T. Porter, R. Stewart, and J. Reed, "Hydration level monitoring using embedded piezoresistive microcantilever sensors," *Med. Eng. Phys.*, vol. 27, no. 3, pp. 215–220, 2005.
- [13] R. Stewart, J. Reed, J. Zhong, K. Morton, and T. L. Porter, "Human hydration level monitoring using embedded piezoresistive microcantilever sensors," *Med. Eng. Phys.*, vol. 29, no. 10, pp. 1084–1088, 2007.
- [14] Y. Y. G. Hoe, B. H. Johari, M. Ju, S. Kim, K. Vaidyanathan, and T. G. Kang, "A microfluidic sensor for human hydration level monitoring," in *Proc. Expo Defense Sci. Res.*, Singapore, Aug. 2011, pp. 1–4.
- [15] F. Scharhag-Rosenberger, "Ergospirometry for endurance performance diagnostics (Spiroergometrie zur Ausdauerleistungsdiagnostik)," *Deutsche Zeitschrift für Sportmedizin*, vol. 61, no. 6, pp. 146–147, 2010.
- [16] B. R. Ely, S. N. Cheuvront, R. W. Kenefick, and M. N. Sawka, "Limitations of salivary osmolality as a marker of hydration status," *Med. Sci. Sports Exercise*, vol. 43, no. 6, pp. 1080–1084, 2011.
- [17] A. C. Guyton and J. E. Hall, *Textbook of Medical Physiology*, 11th ed. Philadelphia, PA, USA: Elsevier, 2006.

- [18] L. E. Armstrong, "Hydration assessment techniques," *Nutrition Rev.*, vol. 63, no. 6-II, pp. S40–S54, 2005.
- [19] S. N. Cheuvront, B. R. Ely, R. W. Kenefick, and M. N. Sawka, "Biological variation and diagnostic accuracy of dehydration assessment markers," *Amer. J. Clin. Nutrition*, vol. 92, no. 3, pp. 565–573, 2010.
- [20] K. Pearson, "Mathematical contributions to the theory of evolution. III. Regression, heredity, and panmixia," *Philosoph. Trans. Roy. Soc. London Ser. A, Containing Papers Math. or Phys. Character*, vol. 187, pp. 253–318, 1896.
- [21] C. Spearman, "The proof and measurement of association between two things," *Amer. J. Psychol.*, vol. 100, nos. 3/4, pp. 441–471, 1987.
- [22] E. J. Wegman, "Hyperdimensional data analysis using parallel coordinates," *J. Amer. Statist. Assoc.*, vol. 85, no. 411, pp. 664–675, 1990.
- [23] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*. Cambridge, MA, USA: MIT Press, 2006.
- [24] S. Faul, G. Gregorčič, G. Boylan, W. Marnane, G. Lightbody, and S. Connolly, "Gaussian process modeling of EEG for the detection of neonatal seizures," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 12, pp. 2151–2162, Dec. 2007.
- [25] O. Stegle, S. Fallert, D. MacKay, and S. Brage, "Gaussian process robust regression for noisy heart rate data," *IEEE Trans. Biomed. Eng.*, vol. 55, no. 9, pp. 2143–2151, Sep. 2008.
- [26] R. Dürichen, M. Pimentel, L. Clifton, A. Schweikard, and D. Clifton, "Multitask Gaussian processes for multivariate physiological time-series analysis," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 1, pp. 314–322, Jan. 2015.
- [27] S. I. Lee *et al.*, "A prediction model for functional outcomes in spinal cord disorder patients using Gaussian process regression," *IEEE J. Biomed. Health Informat.*, vol. 20, no. 1, pp. 91–99, Jan. 2016.
- [28] T. Hastie, R. Tibshirani, and J. H. Friedman, *The Elements of Statistical Learning*, 2nd ed. New York, NY, USA: Springer, 2009.
- [29] V. Vapnik, *The Nature of Statistical Learning Theory*. New York, NY, USA: Springer, 1995.
- [30] B. Schölkopf, A. J. Smola, R. C. Williamson, and P. L. Bartlett, "New support vector algorithms," *Neural Comput.*, vol. 12, no. 5, pp. 1207–1245, 2000.
- [31] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.
- [32] I. H. Witten, E. Frank, and M. A. Hall, *Data Mining: Practical Machine Learning Tools and Techniques*, 3rd ed. Burlington, MA, USA: Morgan Kaufmann, 2011.
- [33] M. Stone, "Cross-validatory choice and assessment of statistical predictions," *J. Roy. Statist. Soc. Ser. B (Methodol.)*, vol. 36, no. 2, pp. 111–147, 1974.
- [34] M. Ring, C. Lohmueller, M. Rauh, J. Mester, and B. M. Eskofier, "A temperature-based bioimpedance correction for water loss estimation during sports," *IEEE J. Biomed. Health Informat.*, to be published.
- [35] V. Wynn, "The osmotic behaviour of the body cells in man; significance of changes of plasma-electrolyte levels in body-fluid disorders," *The Lancet*, vol. 270, no. 7007, pp. 1212–1218, 1957.
- [36] S. N. Cheuvront, R. W. Kenefick, K. J. Sollanek, B. R. Ely, and M. N. Sawka, "Water-deficit equation: Systematic analysis and improvement," *Amer. J. Clin. Nutrition*, vol. 97, no. 1, pp. 79–85, 2013.
- [37] E. T. Perrier *et al.*, "From state to process: Defining hydration," *Obesity Facts*, vol. 7, no. S2, pp. 6–12, 2014.
- [38] S. N. Cheuvront, R. W. Kenefick, N. Charkoudian, and M. N. Sawka, "Physiologic basis for understanding quantitative dehydration assessment," *Amer. J. Clin. Nutrition*, vol. 97, no. 3, pp. 455–462, 2013.

Author' photographs and biographies not available at the time of publication.