Predicting firefighters' exertion based on machine learning techniques

Stefan Kupschick[1], Marie Pendzich[2], Dorata Gardas[1], Thomas Jürgensohn[1], Sascha Wischniewski[2], Lars Adolph[2]

baua: Focus

Firefighters are exposed to many hazards. Smart personal protective equipment can assist in hazardous situations by processing vital parameters. The main challenge in this context is the correct interpretation of individual strain. Therefore individual assessments and physiological measurements were used in a machine learning context. The result is a well-working classification model to estimate the firefighters current exertion.

Content

1 Introduction 1
2 Materials and Method 2
2.1 Participants 2
2.2 Data collection 3
3 Theory 4
3.1 Machine Learning 4
4 Application and Results 6
4.1 Step One: Creating the separating hyperplane 6
4.2 Step Two: Creating a simple and reliable prediction 7
5 Results 8
6 Discussion 10
7 Conclusions 12

1 Introduction

Firefighting is a high risk profession and firefighters are exposed to many hazards. Especially fire suppression with search and rescue tasks in a burning building can cause unpredictable, critical and life-threatening situations. American (U.S.) statistics (Haynes 2015) show that 42,6 % of all firefighter injuries occurred during fire ground operations.

To protect firefighters from burn and inhalation injuries they require personal protective equipment (PPE) including protective clothing and self-contained breathing apparatus (SCBA). The insulation of clothing and the weight of PPE and SCBA represent an additional load for firefighters. Environmental factors such as external heat, extreme workload and a disturbed thermoregulatory system (Montain et al 1994) result in physical exhaustion and thermal stress (Barr et al 2010; Duncan et al 1979). This in turn leads to physiological and

[1]HFC GmbH, [2]Federal Institute for Occupational Safety and Health

baua: Federal Institute for Occupational Safety and Health cardiovascular strain (Louhevaara et al 1995; Smith et al 1997; White et al 1989). In 2015, half of on-duty firefighter deaths in the U.S. resulted from overexertion and related medical issues (Fahy et al 2016). There are some job-related specifics for firefighters: even if they finished fire suppression and leave the source of fire the body core temperature still rises for several minutes (Rossi 2003; Selkirk and McLellan 2004). Another point is that high thermal stress affected the cognitive performance particularly while dealing with complex tasks (Hancock and Vasmatzidis 2003) for example searching for alternative solutions (Perroni et al 2014). These circumstances associated with time pressure - rescue tasks and fire suppression need to be carried out as fast as possible, especially when the SCBA low air alarm sets off - may lead to wrong and hazardous decision-making.

Smart personal protective equipment (PPE) can assist at an early stage in such situations by analyzing, interpreting and visualizing hazards by processing vital and/or environmental parameters. Physiological parameters can be used to predict the health status of firefighters and therefore can support decision-making in critical situations. However, the pure representation as a "Physiological Status Monitoring" is not conducive because it does not enable prediction of overexertion. Individual strain is closely linked to human attributes like gender, age or physical fitness (Smith et al 2008) and interacts with environmental factors with different degrees of intensity. The most important aspect for limiting physical performance in hot environments is the attainment of a critically high body core temperature (González-Alonso et al 1999). The World Health Organisation (1969) recommended that the deep body temperature should not exceed 38 °C for prolonged daily exposures in heavy work. Studies from Montain et al., (1994) White et al. (1989) and Selkirk and McLellan (2004) pointed out that this limit is already exceeded during fire-fighting in protecting clothing at low exercise intensities and normal ambient temperature. Montain et al. (1994) noted "that exhaustion occurred over a broad range of core temperature and that there was no threshold where exhaustion abruptly increased". Similar findings regarding the heart rate were investigated in studies with firefighters from Smith and Petruzzello (1998), Smith et al. (2001), Faff & Tutak (1989). In these studies a great number of subjects reached their age-predicted maximum heart rate or exceeded it partially (Finteis et al 2002) without interrupting the experiment and without consequences for their health. Although extensive research has been carried out on physiological responses of firefighters during fire-fighting simulations, little is known about health-endangering physiological limitations in this special context.

For the efficient use of smart PPE it is therefore necessary to analyze the physiological parameters correctly considering individual human attributes. A new approach is the combination of objective physiological parameters with the subjective perception of effort. Thus, generally accepted limits adjusted to the context of use can be derived. This complex analysis, can be done using machine learning techniques. The basis of this study consists of a comprehensive database, which was generated in field experiments with professional firefighters.

2 Materials and Method

2.1 Participants

39 professional healthy firefighters were tested. However, data from 17 subjects were not included in the analysis because there were technical difficulties (measurement technology of body core temperature) in the fire simulation leading to loss of data. Therefore, data of 22 firefighters was used for machine learning (Table 1).

Table 1 Demographics of firefighters (only subjects with complete data sets)

	n = 22
Age (years)	36,05 ± 7,03
Height (cm)	179,4 ± 7,85
Weight (kg)	80,91 ± 7,37
BMI	25,2 ± 4,4
Firefighter experience (years)	11,14 ± 7,98
Fire suppression training (number of incident)	63,95 ± 67,12
Fire suppression real (number of incident)	22,95 ± 23,99

2.2 Data collection

The data collection was realized in a special test tunnel in which fire exercises can be performed under very realistic conditions. The study was integrated into an existing firefighting training. This had the advantage that working effort and behaviour of the firefighters could be observed unaltered without external influences caused by guidelines and restrictions of standardized tests. The task was to search and rescue fire victims and perform the first fire attack. The type of live-fire, the number of victims and the operation time varied in the trainings. Borg's ratings of perceived exertion scale (RPE; 2004) was used to identify the perceived individual level of exertion during the training. The scale starts at 6 (no exertion) and goes up to 20 (maximal exertion). The RPE-scale was chosen because it is a simple but reliable and a scientifically valid scale to assess the physical strain. Personality and exhaustion can influence self-perception and may lead to over- or underestimation of exertion. To control this and gather semi-objective data, four professional firefighters with many years of experience in fire suppression and training in the test tunnel also evaluated the subjects. In a special workshop the professional firefighters got familiar with the RPE-scale and developed anchor points to assess the physical strain of the subjects consistently. A pre-test was performed to calibrate the assessments.

The measurements consisted of three stages; the first stage served for preparation. The subjects were briefed about the study, the RPE-scale and provided their voluntary and informed consent. Each subject's age, height, weight, gender, smoking behaviour, activity-level (0 = avoid activity to 10 = high endurance athletes), general stress resistance (1 = no resistance to stress to 9 = high resistance to stress) and work experience were obtained. Heart rate was continuously measured by chest belt (Suunto Memory Belt) and the body core temperature was continuously assessed via in-ear-system (cosinuss° GmbH). The subjects wore their own fire-protective clothing which included a multilayer fire-protective suit, fire hood, gloves, helmet, safety boots, SCBA and individual equipment (for example tool belt, hose, infrared camera, fire-axe). After preparing the measurements were started and the physiological parameters were recorded for at least 5 minutes in a relaxed atmosphere. The first internal and external RPE-assessment was reported. The live-fire training was the second stage. The training consisted of approximately 30 min of three simulated firefighting activities: search and rescue fire victims in a special house with limited field of vision, find the "toxic door" to the source of fire and start fire suppression in the tunnel. The ambient temperature rose constantly. During the training the subject was observed and rated permanently by an expert in 5-minutes intervals. This rating, including a detailed description of the current situation, was transmitted to investigators outside of the tunnel. These notes served as a reminder and enabled the subjects to evaluate the same moments like the external raters. The self-assessment was carried out retrospectively after training, so that the subjects could fulfill their task/

training without any interruption and limitation of safety. Upon completion of the training an After-Action-Review with the trainer followed. The subjects rested for 15 minutes and the investigators asked for their assessment of perceived exertion during the live-fire training. The internal and external assessment continued in 5-minutes intervals. At the end the National Aeronautical and Space Administration Task Load Index (NASA-TLX; Nienaber 1997) was completed by the subjects and the external raters. The NASA-TLX was not used in the machine learning context. In total data of 19 parameters were measured and four questionnaires were fulfilled (Table 2).

Table 2 Overwiev collected data

Demografic parameters	
Age (years)	
Height (cm)	
Weight (kg)	
BMI	
Firefighter experience (years)	
Fire suppression training (number of incident)	
Fire suppression real (number of incident)	
Lung volumen	
Physiological parameters, measured	**derived parameters**
Body core temperture [°C]	temperature slope [°C/t]
Heart rate [bpm]	EPOC [ml/kg]
	Ventilation [l/min]
	VO2 [ml/kg/min] Energy [kcal] breathing rate [bpm] maximum heart rate [bpm]
Questionnaire	
Borg's RPE-Scale	
NASA-TLX	
General resistance to stress	
Activity index (Fletcher 2008)	

3 Theory

The physiological measurements as well as the internal (self-assessment) and external ratings were then used within a machine learning context. With these data sets support-vector-machine algorithms were tested in combination with greedy search procedures to analyse their prediction accuracy.

3.1 Machine Learning

Machine Learning is generally speaking each and every generation of knowledge from experience. In this specific context knowledge consists of functions respectively class assignments whether experience describes structured data. In this case we wanted to train an algorithm to separate two datasets, whose classes were already known, this is called supervised learning.

The procedure can be briefly described as followed: The feature vectors x and their labels y from the result space Y describe the training data, from which a classifier can generate a model. With this model one can assign a label y' to a new feature vector x' (Figure 1).

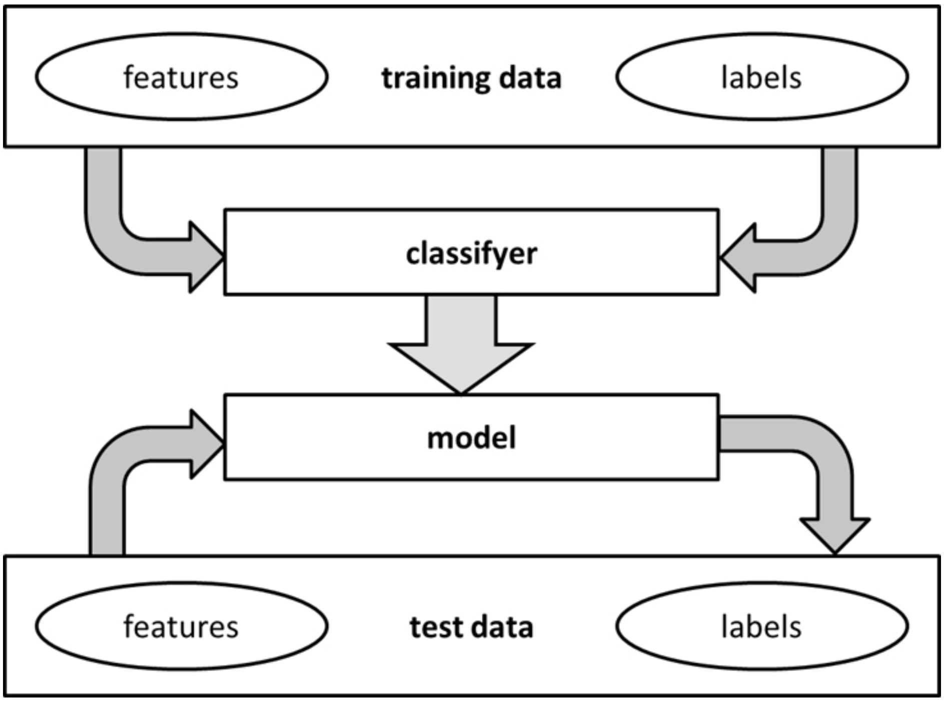

Fig. 1 Supervised machine learning

3.1.1 Support Vector Machine

In this study a support vector machine (SVM) classification algorithm (Vapnik 1995) was applied. A SVM constructs a multidimensional hyperplane into the feature space, which separates the classes from each other (Figure 2). SVM classification can - unlike other algorithms - solve non-linear problems by transforming the input data into high-dimensional feature spaces and then applying linear separation. Therefore different kernel functions can be used to accomplish this transformation, usually RBF (Radial Basis Function) -kernels, as also in this study.

3.1.2 Feature Selection

Because the training data may contain features, which do not add any useable information to the current feature set or even ones, which do not contain any useable information at all a feature selection method should be applied. Different approaches can be used to reduce the number of features. On the one hand "wrapping" techniques and on the other hand so called "filter" techniques. The wrapper-approach applies the whole classification algorithm on a specific subset of features, which is changed according to the feature selection algorithm until an optimal subset is found. The "filter"-techniques directly use the information content of the features, which is the much faster method, but because of no existing interaction between the feature selection process and the classification process one cannot derive any conclusions referring to the classification problem.

Three wrapper algorithms have been implemented, "Sequential Forward Selection" (SFS), "Sequential Backward Selection" (SBS) and "Plus-L-Take-Away-R" (PTA), which is a combination of the first two algorithms All of these procedures base on the same rule: find a feature set F_m with m features and $m < n$, which minimizes the classification error, with n standing for the number of features before feature selection. An exhaustive search would need to evaluate n over m features for n - this is even at low numbers of n not a feasible solution.

SFS starts with an empty set of features and adds than the feature that minimizes the classification error, when combined with the already selected features. SBS starts with a full feature set and removes the one, which contributes least to the minimization of the classification error. The PTA-algorithm combines the SFS and SBS approaches: after adding L features to the subset (with SFS) R features are removed (with SBS). This is repeated until the desired number of features is achieved. The advantage of this slightly slower algorithm is to avoid getting stuck in a local maximum.

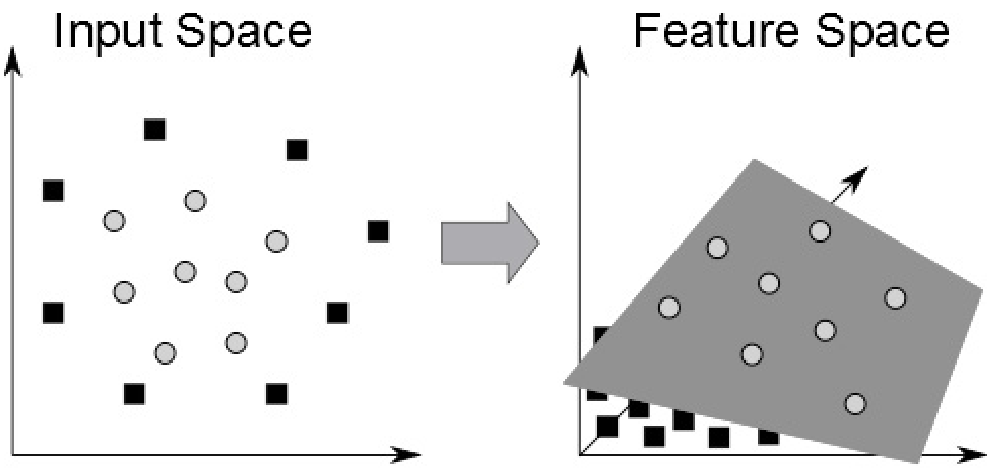

Fig. 2 Mapping from input data to feature space using a kernel function - complex problems in lower dimensions can become simple problems in higher dimensions

4 Application and Results

The method for the generation of an algorithm for the prediction of firefighter's strain consists of two steps: At first an optimal separation between the two classes "high strain" and "low strain" is computed using the before mentioned machine learning and feature selection techniques and a function was computed to fit the separating hyperplane created by the classifier. In a second step a new measure is introduced, which is based on the distances between the particular data points and the approximated separating hyperplane.

4.1 Step One: Creating the separating hyperplane

After visual and algorithmic verification of the recorded data, pre-processing and synchronization were conducted, resulting in two databases available for classification: 331 combined datasets (demographic, physiological including temperature data) and 564 datasets without temperature data. Then, the feature selection methods have been carried out in combination with the before mentioned SVM classification. All classification algorithms used Leave-One-Out cross validation to prevent over-fitting.

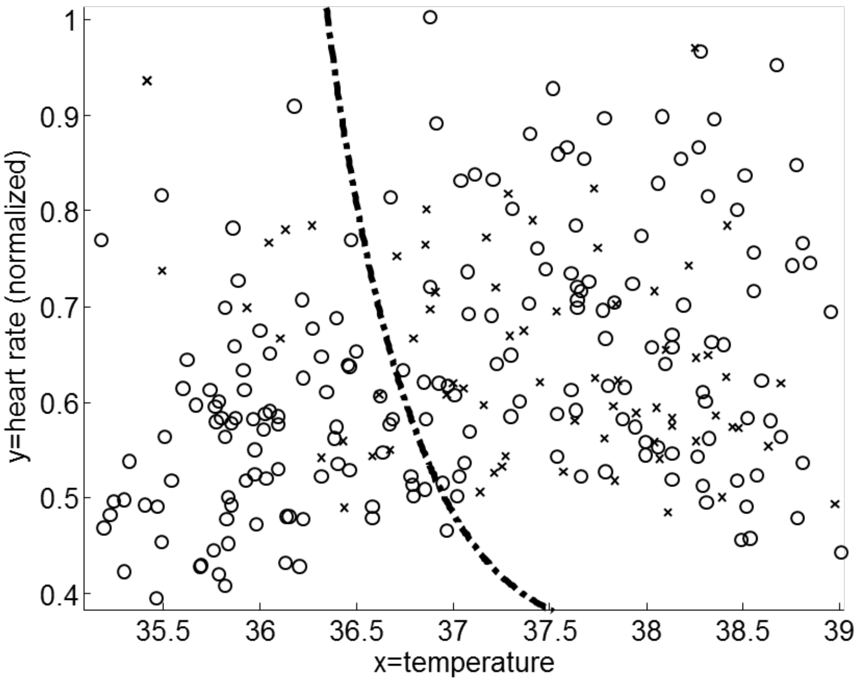

Fig. 3 Two-dimensional approximation function, circles are correct classifications, crosses are misclassified data points

To conduct a 2-class-classification, the 15 point RPE-scale had to be transformed into a 2-point scale. To achieve this, the ratings below '12' were categorized into class 'A', whether all ratings above '13' were grouped to class 'B'. Class A therefore stands for "low strain" and class B for "high strain".

Feature selection algorithms were used to obtain the most relevant features. Results show, that, using more than three different parameters does not lead to a significant increase of the classification rate. Therefore, different separating hyperplanes were computed, dependent on the particular dataset.

4.2 Step Two: Creating a simple and reliable prediction

The multidimensional hyperplanes were created by the SVM-algorithm using nonlinear transformations and kernel functions. Hence, they cannot be represented by simple mathematical functions. To simplify such a classification it is practical to use a polynomial approximation of such a hyperplane. An example of an approximation function is presented in Figure 3. As reflected in the data distribution, the number of misclassifications decreases with ratings in the upper and lower range of the scale: the ratings '12' and '13' - belonging to the classes 'A' and 'B' - are the classes with the most misclassifications. Their data points in feature space are those closest to the separating hyperplane, respectively the approximation function. A simple, but very effective solution to prevent a high number of these false alarms/misses would be to introduce a new measure of information value of a classification. This measure, introduced as "score" is the Euclidean distance between a data point and the approximated hyperplane and provides the possibility to get more information about the validity of the classification of this specific data point. Thus, a larger distance to the hyperplane implies a more reliable classification.

Secondly, a function can be developed, which predicts not only the class membership of a data point, but the rating itself by using the regression function of the before mentioned

score and the particular ratings. Assuming a distance function d $(\tilde{X})$ which assigns a distance between a measurement vector $\tilde{X}$ to the separating hyperplane H constructed by the SVM-algorithm. One can make a connection between $d(\tilde{X})$ and the original BORG-ratings Y. This regression function is depicted in figure 4. For classification using two-dimensional feature vectors consisting of heart rate and temperature rate,
$d(\tilde{X}) = 0.078 * Y - 0.85$ mathematically describes the before mentioned connection. With this knowledge of the correlation between BORG-rating and distance one can easily get a formula for the estimation of Y by rearranging the equation. This leads to $Y' = (d(\tilde{X}) + 0.85) / 0.078$. Assuming a data point $\tilde{X}$ consisting of the measured body core temperature of 38.6 °C and a standardized heart rate of 0.8 obtains a score of 0.493, which - included in to the formula - leads to an estimated Borg-rating Y' of 17.345.

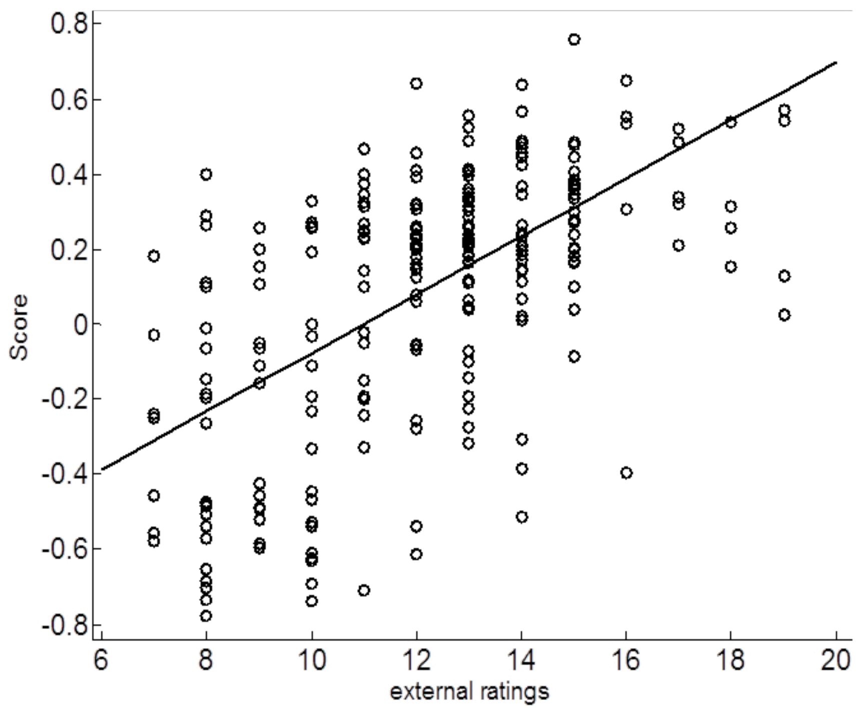

Fig. 4 Mean score in relation to external rating

5 Results

Before proceeding with further computations, the optimal separating hyperplane has to be chosen. In this chapter, a summary of these computations is presented. To obtain the optimal separating hyperplane, different classification computations have been conducted. The following parameters have been varied: The number of features, the rating type (self-assessments vs. external ratings), the feature extraction algorithm (SFS vs. PTA), the dataset (datasets with or without temperature data) and the "gap size". The parameter “gap size” is introduced to cope with the problem of classifying medium Borg-ratings. For example, "gap size=4" stands for removing four ratings, e. g. the ratings '11' - '14' from the database. Therefore, using "gap size=4", class A consists of all datasets with ratings of '6'-'10' and class B consists of all datasets with ratings of '15'-'20'.

Table 3 Number of datasets

Data sets with	External Ratings	Self-Assessments
overall	623	608
physiology including temperature	276 (SET A)	287 (SET B)
physiology without temperature	533 (SET C)	517 (SET D)

Table 3 shows the number of recorded datasets in relation to the corresponding ratings and the type of the recorded data. After rejecting incomplete datasets, 533 respectively 517 datasets with physiology remained. 276 complete datasets with external ratings respectively 287 datasets with self-assessments using physiology and temperature sensors could be obtained, so different computations have been conducted for datasets with and without temperature data.

Table 4 Effect of the number of features, rating type and data type (PTA, all data, in %)

Number of features	1	2	3	4
External ratings (Set A)	73,2	76	77,3	78,3
Self Assessments (Set B)	72,6	73,3	73,9	74,6
External ratings (Set C)	63,6	64,4	67,2	67,2
Self Assessments (Set A)	64,5	66	66	66

Table 4 shows some summarized results: Computations using the PTA-algorithm as feature selection method and using all available ratings (gap size = 0) are compared to show the effect of the number of features used by the classification algorithm. The four datasets described in Table 1 are used and lead to similar results: The difference between using one feature and using five features is only about 1.5 - 5 %. Using datasets without temperature data decreases the number of features, which are needed to reach the best possible classification result. If the algorithm is forced to use only one feature, "temperature" is chosen first, if available. For a database without temperature data, "heart rate" would be the most effective single feature. These findings do not differ between self-assessments and external ratings. The use of temperature data combined with the other data leads to increases of the classification rate of 8 - 12 %. So, measurement of body core temperature data is recommended. Also the use of external ratings instead of self-assessments leads to an increase of classification rates of about 1 - 4 %.

Table 5 Effect of the class gap, rating type and data type (PTA, 5 parameters)

Number of features	0	#	2	#	4	#	6	#
External ratings (Set A)	78,3	276	88,5	190	92,4	140	100	88
Self Assessments (Set B)	72,8	287	74,6	182	84,4	120	97,4	70
External ratings (Set C)	67,2	533	74	381	81,2	290	89,2	194
Self Assessments (Set A)	66	517	77,1	340	83,5	241	87,4	156

Table 5 clarifies the effect of removing specific datasets from the database. "Gap size" describes the number of ratings/datasets which are removed. This is implemented to obtain a "smoother" separating hyperplane: By forcing the SVM-algorithm to not use data vectors near the class border, the probability of over-fitting is highly reduced. Therefore much easier to approximate hyperplane functions are received from the SVM-algorithm. Because of removing data affects the size of the dataset, the number of data used is denoted in brackets behind the percental values. The effect of these removals amounts up to 35 % classification rate and shows clearly, that more extreme ratings lead to drastically increased classification rates.

Table 6 Parameter chosen by the PTA-algorithm using all data (Temp - body core temperature, HR - heart rate, dTemp - increase of body core temperature, BMI - Body Mass Index, BR - breathing rate, Vent - ventilation, VO2 - oxygen uptake

Number of features	1	2	3	4
External ratings (Set A)	Temp	HR	dTemp	BMI
Internal ratings (Set B)	Temp	Vent	HR	BR
External ratings (Set C)	HR	EPOG	Vent	-
Internal ratings (Set A)	HR	VO2	-	-

Table 6 shows, which parameter was chosen in which computation cycle by the PTA-algorithm. In the first cycle body core temperature was chosen, if available. If not, HR was chosen first, followed by diverse heart rate - derived parameters. An exception is the BMI, which was chosen as fourth best parameter using external ratings and Set A and is a constant parameter pertaining to a specific subject during an experiment. Considering the comparatively small increase of the classification rate after the first parameters, the order and selection of the following parameters does not seem to have a large informative value. Besides, the utilized parameters except Temperature and Heart Rate are derived parameters from the aforementioned ones.

6 Discussion

The aim of this study was to subsume physiological parameters under one individual strain parameter, which predicts the rate of exertion under extreme conditions. This was accomplished by using demographic, physiological and temperature data as well as self-assessments and external ratings of perceived exertion. Various datasets were selected to represent the whole range of individual attributes, which interact with environmental factors. Machine learning techniques combined the complex datasets with the subjective perception of effort and derived limits adjusted to the context of use.

The most important result of this study is that machine learning techniques can be applied to identify high and low individual strain reliably. The possibility to simplify the created multidimensional hyperplanes with the help of polynomial approximation forms the basis to use this kind of technology in real fire operations. It had been proven that already simple mathematical calculations with two physiological parameters lead up to 85,8 % of correct classifications in an adapted setting (Table 7). This is significant because it minimizes the complexity of the algorithm and enables the classification of exertion on a mobile device integrated in smart-PPE.

Table 7 Accuracy of classification: Body core temperature combined with heart rate divided into low and high strain with gap size = 4

External rating	Number of datasets	Correct classification	Accuracy
Low strain Borg scale: 6 - 10	90	74	82,2 %
High strain Borg scale: 15 - 20	79	71	89,8 %
Total number of sets	169	145	85,8 %

A further step towards implementation of the findings in real fire operations is the correlation between Borg-rating and Euclidean distance. The score increases proportionately to the value of external ratings and not only allows the prediction of exertion but also provides the opportunity of visualizing the result clearly. This is achieved by dividing the derived regression function into three sections. Each section stand for a status of exertion and forms a kind of traffic light labelling in which the red light stands for maximal exertion (Figure 5). This representation enables a quick overview, for example the fire commander and supported rapid decision making.

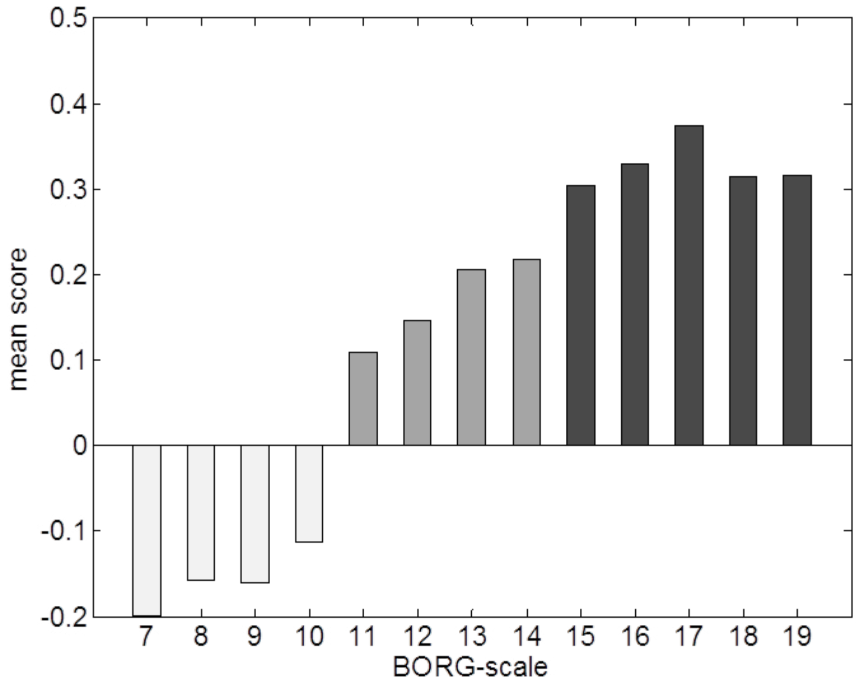

Fig. 5 Traffic light labelling for three classes of strain - class "green" (score <-0.1), class "yellow" (-0.1 ≤ score < 0.3) and class "red" (score ≥ 0.3)

In how far the findings can be applied in real fire operations has to be verified. There are two approaches: (1) check the validity of the computed algorithms under laboratory conditions and (2) generate new training data in real-fire operations and compare them with the data from the fire-fighting simulation.

The goals of the laboratory testing are on the one hand the transferability potential of the project results to other professionals groups (e. g. occupational divers) and on the other hand the clarification of mismatch between external ratings and self-assessments of perceived exertion. Real-fire operations are defined by a sudden increase in physical activity (Barr et al., 2010) and a high level of psychological stress (Moran & Colless, 1995), for example when internal and external conflicts of the individual about goals and consequences of action occur (Luczak, 1991). Firefighting simulations do not sufficiently reflect this stress-strain relationship. It is necessary to generate some training data during real-fire operations. The aim is to achieve more ratings in the range between 17 and 20 on the Borg-scale to approximate to physiological limits.

7 Conclusions

The results show a promising approach to subsume physiological parameters under one individual strain parameter (rate of exertion) and to visualize it. Therefore it is possible to identify high and low individual strain reliably using complex data. This can help to use health-monitoring smart personal protective equipment to increase safety at first for firefighters - other fields of application are conceivable. Open questions are whether the identified algorithm is generally valid, how to cope with the existing analytical imprecision as well as legal and behavioural aspects. Research is conducted on these issues in other projects of the Federal Institute for Occupational Safety and Health.

References

[1] Barr, D., Gregson, W., Reilly, T. 2010. "The thermal ergonomics of firefighting reviewed." Applied Ergonomics 41: 161-72

[2] Belding, H.S., Givoni, B., Gupta, M.N., Lavenne, F., Lind, A.R., Metz, B., Shahbazjan, S.H., Wenzel, H.G. 1969. "Health Factors involved in working under Conditions of Heat Stress."

[3] Borg, G. 2004. "Anstregungsempfinden und körperliche Aktivität." Deutsches Ärzteblatt 15: 1016-21

[4] Duncan, H.W., Gardner, G.W., Barnard, R.J. 1979. "Physiological Response of Men Working in Fire Fighting Equipment in the heat." Ergonomics 22(5): 521-27

[5] Faff, J., Tutak, T. 1989. "Physiological responses to working with fire fighting equipment in the heat in relation to subjective fatigue." Ergonomics 32(6): 629-38

[6] Barr, D., Gregson, W., Reilly, T. 2010. "The thermal ergonomics of firefighting reviewed." Applied Ergonomics 41: 161-72

[7] Belding, H.S., Givoni, B., Gupta, M.N., Lavenne, F., Lind, A.R., Metz, B., Shahbazjan, S.H., Wenzel, H.G. 1969. "Health Factors involved in working under Conditions of Heat Stress."

[8] Borg, G. 2004. "Anstregungsempfinden und körperliche Aktivität." Deutsches Ärzteblatt 15: 1016-21

[9] Duncan, H.W., Gardner, G.W., Barnard, R.J. 1979. "Physiological Response of Men Working in Fire Fighting Equipment in the heat." Ergonomics 22(5): 521-27

[10] Faff, J., Tutak, T. 1989. "Physiological responses to working with fire fighting equipment in the heat in relation to subjective fatigue." Ergonomics 32(6): 629-38

[11] Fahy, Rita F., LeBlanc, Paul R., Molis, Joseph L. 2016. "Firefighter Fatalities in the United States-2015." National Fire Protection Association.

[12] Finteis, T., Oehler, J.-C., Genzwürker, H., Hinkelbein, J., Dempfle, C.-E., Becker, H., Ellinger, K. 2002. "Stressbelastung von Atemschutzgeräteträgern bei der Einsatzsimulation im Feuerwehr-Übungshaus Bruchsal Landesfeuerschule Baden-Württemberg."

[13] Fletcher, Eddie. 2008. "Suunto t6c running guide."

[14] González-Alonso, J., Teller, C., Andersen, S.L., Jensen, F.B., Hyldig, T., Nielsen, B. 1999. "Influence of body temperature on the development of fatigue during prolonged exercise in the heat." Journal of Applied Physiology 86(3): 1032-39

[15] Hancock, P.A., Vasmatzidis, I. 2003. "Effects of heat stress on cognitive performance: the current state of knowledge." Int.J.Hyperthermia 19(3): 355-72

[16] Haynes, H.; Molis, J. 2015. "U.S. Firefighter Injuries - 2014." National Fire Protection Association.

[17] Louhevaara, V., Ilmarinen, R., Griefahn, B., Künemund, C., Mäkinen, H. 1995. "Maximal physical work performance with European standard based fire-protective clothing system and equipment in relation to individual characteristics." European Journal of Applied Physiology 71: 223-29

[18] Montain, S.J., Sawka, M.N., Cadarette, B.S., Quigly, M.D., Mckay, J.M. 1994. "Physiological tolerance to uncompensable heat stress: effects of exercise intensity, protective clothing, and climate." Journal of Applied Physiology 77(1): 216-22

[19] Nienaber, C. 1997. "Psychische Beanspruchung im Assessment Center." Academy of Management Journal.

[20] Perroni, F., Guidetti, L., Cignitti, L., Baldari, C. 2014. "Psychophysiological Response of Firefighters to Emergencies: A Review." The Open Sports Science Journal 7: 8-15

[21] Rossi, R. 2003. "Fire fighting and its influence on the body." Ergonomics 46(10): 1017-33

[22] Selkirk, G.A., McLellan, T.M. 2004. "Physical Work Limits for Toronto Firefighters in Warm Environments." Journal of Occupational and Environmental Hygiene 1(4): 199-212

[23] Smith, D.L., Horn, G., Goldstein, E., Petruzzello, S.J. 2008. "Firefighter Fatalities and Injuries: the Role of Heat Stress and PPE."

[24] Smith, D.L., Manning, T.S., Petruzzello, S.J. 2001. "Effect of strenuous live-fire drills on cardiovascular and psychologigal responses of recruit firefighters." Ergonomics 44(3): 244-54

[25] Smith, D.L., Petruzzello, S.J. 1998. "Selected physiological and psychological responses to live-fire drills in different configurations of firefighting gear." Ergonomics 41(8): 1141-54

[26] Smith, D.L., Petruzzello, S.J., Kramer, J.M., Misner, J.E. 1997. "The effects of different thermal environments on the physiological and psychological responses of firefighters to a training drill." Ergonomics 40(4): 500-10

[27] Vapnik, Vladimir. 1995. "The Nature of Statistical Learning Theory." Springer.

[28] White, M.K., Vercruyssen, M., Hodous, T.K. 1989. "Work toleranceand subjective response to wearing protective clothing and respirators during physical work." Ergonomics 32(9): 1111-2

Kupschick, S.; Pendzich, M.; Gardas, D.; Jürgensohn, T.; Wischniewski, S.; Adolph, L. (2016). Predicting firefighters‘ exertion based on machine learning techniques. Dortmund: Bundesanstalt für Arbeitsschutz und Arbeitsmedizin (baua: Focus). doi:10.21934/baua:focus20161107

Human Factors Consult

Imprint | Publisher: Federal Institute of Occupational Safety and Health (BAuA), Friedrich-Henkel-Weg 1-25, 44149 Dortmund, Germany | Contact: Marie Pendzich, Phone: +49 231 9071-2150, E-Mail: *pendzich.marie@baua.bund.de*, Internet: *www.baua.de*, Design: eckedesign Berlin, R. Grahl, (BAuA) | DOI: 10.21934/baua:focus20161107 | November 2016

baua: Federal Institute for Occupational Safety and Health