Towards Decisive Garments for Heat Stress Risk Detection

Mitra Baratchi
University of Twente

Lennart Teeunissen, Peter Ebben, Wouter Teeuw
Saxion University of Applied Sciences

Jan Laarhuis
Thales Nederlands

Maarten van Steen
University of Twente

ABSTRACT
One of the numerous applications of wearable computers is providing safety in occupations where heat-related injuries are prevalent. Core temperature, as a measurement that cannot be measured by on-body sensors is a variable that is specifically interesting for realizing such applications. In the context of the design of a sensor-shirt that can be used for firefighters, in this paper we study the importance of different types of sensor measurements and their placement for predicting core temperature. We propose a model for inferring the dangerous states of core temperature. Our evaluation results show that our model can to a great extent predict hazardous situations caused by heat accumulation.

Author Keywords
Wearable computing; data mining; machine learning; ambulatory physiological monitoring.

ACM Classification Keywords
H.5.m. Application-based systems : Miscellaneous

INTRODUCTION
Heat stress may cause a person's body temperature to rise above the hyperthermic threshold of 37.8-38.3°C. Prolonged and/or severe **hyperthermia** may result in disability or even mortality. This situation is a common casualty among people with occupational heat exposure such as firefighters, mineworkers, and any other type of profession which involves operations in high temperatures, radiant heat sources, high humidity, direct physical contact with hot objects, protective clothing, or strenuous physical activities.

Heat stress is directly detectable by measuring a person's **core body temperature**. Common ways of measuring core temperature are, however, oftentimes invasive as in swallowing core temperature pills (intestinal), or using rectal probes. Such measurement methods are not practical for use on an occupational basis. For instance, temperature pills need to be swallowed at least two hours before measurements are valid and restrict the user to go into an electromagnetic field as long as the pill resides inside the body. Having the possibility of measuring many types of physiological parameters, wearable sensors may provide the potential of estimating core temperature from other measurements. However, it remains an open question which sensors can or should be used for such measurements, and how measurements are to be combined to come to accurate estimations. Difficulties are aggravated when requiring that sensing should take place in an unintrusive fashion.

In this paper, we explore accurate estimation of core body temperature through highly unintrusive wearable sensors. In particular, we concentrate on developing a model that takes an existing dataset containing input from multiple on-body and environmental sensors and ask ourselves: (1) which sensor measurements are important for accurate temperature estimations, and (2) how machine learning techniques can be used to create a model which can capture the hazards of heat exposure. While previous research has focused on accurate estimation of core temperature, our main contribution is showing that although accurate estimation of core temperature may not be easily feasible, a limited number of well-placed on-body sensors can be effectively used to obtain an accurate classification of hazardous and non-hazardous core temperatures. This is crucial for developing specific garments with built-in sensors for people operating in hazardous environments, such as firefighters.

We organize the rest of this paper as follows. In the following section we will first introduce the related work. In Section 3 we discuss our research methodology in developing a model. In Section 4, we subsequently evaluate our proposed solution. We come to conclusions in Section 5.

RELATED WORK
The application of wearable technologies targeting the risks of occupational heat exposure has recently gained attention. There are in general two groups of researchers who have focused on dealing with such risks. The first group has focused on designing wearables that can measure physiological parameters that are relevant for heat-related injuries. The focus of such research is in providing a wearable infrastructure that can measure and transmit certain biometrics without further analysis of the data [5, 8, 9].

The other group of researchers have focused on the important problem of estimating core temperature and physiological strain. For instance, derivative body temperatures have been studied as an estimate for core temperature [10]. In addition, indicators based on internal and external factors (physiological strain and heat stress indices) have been proposed to estimate the thermal load on the body [3, 7]. All in all, the results of this research generally suggests that correct estimation of core temperature (especially, in a non-controlled setting) is a challenging problem [11]. When such estimation is needed to be done from rawearables, the challenge becomes even bigger. To the best of our knowledge, there is little research proposing solutions for estimating core temperature from non-invasive physiological parameters provided by wearable computers [2,4]. The work most in-line with ours is the one proposed in [4]. The authors have proposed a model for estimating heat-stress risk without considering the importance of each of these sensor measurements. They use the mean skin temperature on various spots as a proxy for core temperature. Nevertheless, none of the previous studies have considered the use of different non-invasive sensors, considering both the operational restrictions, and their placement, as well as, an accurate inference model. These are all essential requirements for designing a sensor-shirt to be used in critical situations.

RESEARCH METHOD

Our main goal is to predict the occurrence of health risks due to heat stress. As core temperature is not easy to measure directly with sensors, we study the possibility of predicting it through sensors embedded in a shirt. A standard approach for doing so consists of two steps. **1) Feature selection:** to study the importance of a set of sensor measurements and their placement in correct estimation of the core temperature. **2) Model development based on the selected measurements:** to correctly predict core temperature. In this case, a number of algorithms are used to create a model and the algorithm which provides better results in terms of different performance metrics is chosen. In order to perform the above-mentioned, we use an available multi-parameter dataset provided by [1] which is collected during a firefighting training. This dataset contains time-series data of 12 male subjects who have performed different activities both in normal (20 °C) and extreme thermal conditions (40 °C). Several parameters are present in the dataset such as heart rate (HR), temperature in different body parts, rectal core temperature, phase change material (PCmaterial), activity type, and the cooling type.

Feature selection: What sensors to choose from?

By performing feature selection on the dataset, we will be able to select the minimal-optimal model. This means that we can decide on the placement of sensors on the shirt such that we have a better estimate of core temperature but also such that the choice of features (which leads to the choice of sensors) is practical with respect to the design of a sensor-shirt. Before feature selection, a large number of outliers, such as core and aural temperatures below 35 °C, are removed from the dataset.

Looking at the correlation coefficients, presented in Figure 1, we see that the aural temperature is the temperature which is mostly correlated with core temperature followed by heart rate, thigh and calf temperatures. The correlation coefficients in this manner only represent the linear relationship between two variables. Therefore, we also rank features using a recursive feature selection algorithm in terms of their importance in the prediction of core temperature. Using a feature selection algorithm, many models are trained with different subsets of the model and the best set of features is chosen. The overall of features is presented in Figure 2. For this purpose we have used an implementation based on Random Forest Algorithm which adds random features and remodels [6]. As seen in Figure 2, none of the features have been found as unimportant. The most important feature within these features in prediction of core temperature are the candidate's identity, and aural temperature. The duration of activity (Deltat) appears to be the least important. Studying the results of the above feature ranking, looking at the correlation coefficients, and considering the practical consideration of designing a sensor shirt, we select the partial set of features composed of; trial type, PCmaterial temperature, activity, chest temperature, and HR. The most important features in the prediction of core temperature which are aural temperature and the candidates identity are not considered. The reason is that the aural temperature cannot be practically measured using a shirt. Considering the candidate's identity will lead to develping personalized models for heat stress which is desirable in terms of accuracy. However, we ignore this option as it implies training the model for each individual who is going to use a sensor shirt.

The result of the prediction with a complete and partial set of features using a random forest algorithm is presented in Figure 3. We see that there is a considerable difference between predictions with the complete set of features and the partial set. Nevertheless, in the following section we show how this partial set can also be used in creating an accurate model for avoiding the hazardous situations.

Model development: Why predicting core temperature is not an option?

Having knowledge about core temperature, the heat-stress risk can be estimated. Therefore, the most obvious solution in estimating heat stress risk is correct estimation of core temperature. In this section, we show why such an approach does not work. In order to test the viability of this idea we formulate a problem to estimate core temperature from a set of other measurable features and we compare different algorithms in solving this problem. The only requirement that these algorithms need to satisfy is having the possibility of predicting a numeric output (rather than a categorical output). In other words, they should be regression-based rather than classfication-based. Among many different possible algorithms, we have chosen the following:

Bayesian networks, which are well-known models for allowing domain expert knowledge as input. To use such input as well, we chose this algorithm. We used the hill climbing algorithm to learn the model structure and the parameters of the continuous variables are learnt by Gaussian distribution on the discovered model structure. Domain knowledge was further used to refine the model.

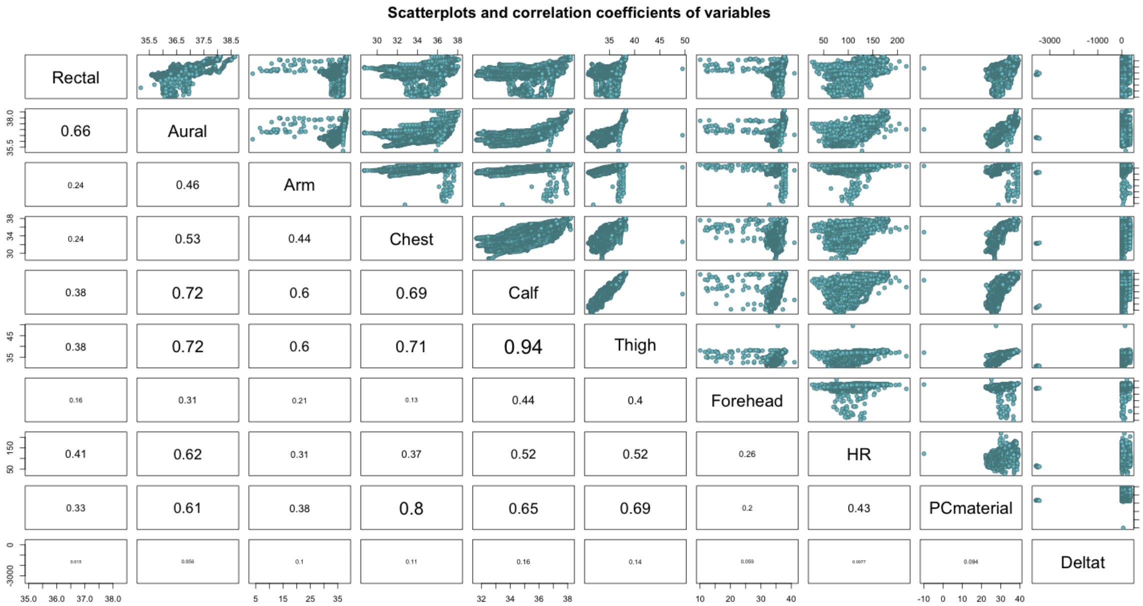


Figure 1: Scatter plots and correlation coefficients of the features.

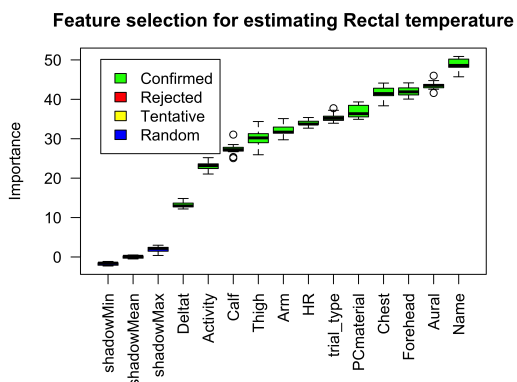


Figure 2: Ranking of different features in prediction of core temperature.

Neural networks, which are less intuitive for domain experts to understand but sometimes lead to better models than Bayesian networks.

Multi-linear regression, which is strong in capturing linear relationships. As seen in the scatterplots before, there is a considerable linear relationship between variables. Therefore, we also use linear regression to predict the core temperature from the partial set of variables.

We compare the performance of these algorithms in terms of the error metrics mentioned below which are used for evaluating regression problems. In what follows, a_t represents the actual value and f_t the forecasted value, respectively:

Root Mean Squared Error (RMSE), provides the sample standard deviation of the difference between the predicted and observed values. $RMSE = \sqrt{\sum_{t=1}^{n}(f_t - a_t)^2/n}$

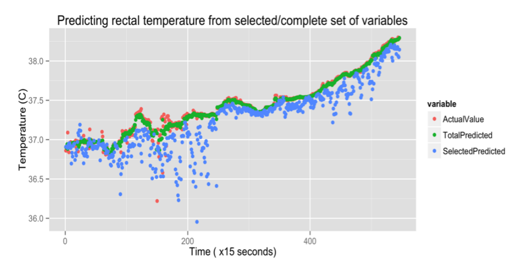


Figure 3: Prediction of core temperature form complete and partial set of features.

Mean Absolute Error (MAE), shows how close forecasts are to the eventual outcome. As this metric is measured in the same unit as the data, it is more understandable than the rest of metrics. $MAE = 1/n\sum_{t=1}^{n}(|f_t - a_t|)$

Mean Error (ME), also measures the difference between the predicted and actual value in the same unit as the original data. However, being disproportionately positive or negative it is more robust to bias in the forecasts. This way models that do consistently under/over estimate are found.

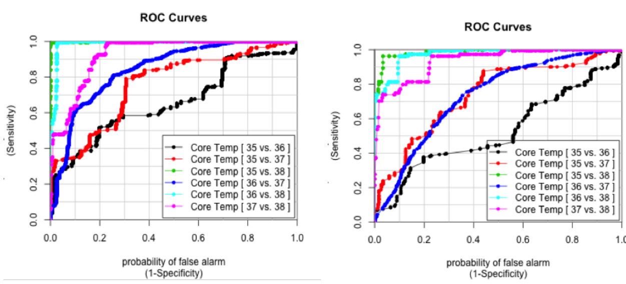
Mean Absolute Percentage Error (MAPE), is the average of the magnitude of error with respect to the magnitude of the actual value. $MAPE = 1/n\sum_{t=1}^{n}(|(f_t - a_t)/(a_t)|)$

Mean Percentage Error (MPE), is an average of the percentage errors with respect to the magnitude of the actual value. $MPE = (100\%/n)\sum_{t=1}^{n}((a_t - f_t)/(a_t))$.

Table 1 compares the accuracy of the previously mentioned algorithms in terms of these error metrics. None of the algorithms is consistently better than the rest. As seen, the RMSE of these algorithms which shows the average temperature er-

ALgorithms	ME	RMSE	MAE	MPE	MAPE
Regression	0.003	0.4250	0.2858	0.0049	0.775
Bayesnet	0.0007	0.5336	0.3695	0.0152	0.999
NNet	0.0067	0.4101	0.2757	0.0177	0.744

Table 1: Comparison of different error metrics in forecasting core temperature from selected features.



(a) (b)

Figure 4: ROC curves for comparing the classification accuracy of hazardous/safe classes from (a) complete (b) partial set of features.

ror, is in range of 0.4 $\sim$ 0.5 °C. From the medical point of view, this is a rather large error value and mostly unacceptable for critical medical applications. That is why building a model for accurately predicting the core temperature is not a useful option. The ultimate goal for predicting core temperature is to infer only the hazardous events that will lead to heat stress. Therefore, it is important to only correctly capture the nonhazardous core temperatures from the hazardous ones. Prediction accuracy of lower temperatures does not provide any advantage in an application such as firefighting. A model that can only predict temperature passing the danger threshold and distinguish it from the temperatures below the danger threshold will meet the requirements.

Model developement: Heat stress hazard model as an alternative to core temperature prediction

Our important requirement is that temperatures in the heat stress zone are correctly distinguished from those that are not. Therefore, apart from the previous models that treat core temperature as a numerical variable we perform analysis on core temperature by treating it as a two-class categorical variable; dangerous (above 38°C)/safe (below 38°C). In other words, we turn the previous regression problem into a classification problem. To build the new model, we discretized the core temperature to its integer unit range. Next, we built and learnt a Bayesian network on the discretized data.

EVALUATION

As this is a classification problem, we cannot use the previously mentioned error metrics for evaluation. ROC curves can better be used for evaluation of a classification model through cross-validation. In such curves, a bigger area under the curve would suggest better classification. We have trained the Bayesian network both from the complete and partial set of features resulted from the feature selection phase. ROC curves and the prediction accuracies are presented in Figure 4 and Table 2, respectively. Although using the complete set

Class	Selected features	All features
35 vs. 36	0.52	0.65
35 vs. 37	0.73	0.75
35 vs. 38	0.98	0.99
36 vs. 37	0.73	0.84
36 vs. 38	0.97	0.98
37 vs. 38	0.93	0.92

Table 2: Classification accuracy of different core-temperature classes using complete and partial set of features.

of features in general might result in a higher area under the curve, the partial set of features can also perform well, especially in distinguishing dangerous temperatures from non-dangerous ones. In other words, in distinguishing the ones above 38°C from the ones below (35 vs 38, 36 vs 38, 37 vs 38).

Prediction accuracies presented in Table 2, also confirm that in general by using all features we can have higher accuracy for both temperatures in hazardous and non-hazardous zones. However, the same plausible results can be achieved from the partial set of features for only the temperatures above 38°C.

CONCLUSION

In this paper, we proposed a model for real-time analysis of physiological parameters for estimation of heat stress risk. Using a dataset collected during a firefightling training session, we studied the importance of different types of sensors, their placement and effectiveness in prediction of core temperature. Our results show that although precise prediction of core temperature with high accuracy is a challenge, it is still possible to classify core temperature in hazardous and safe zones with high accuracy.

REFERENCES

1. Brown, F. Physiological, perceptual and cognitive responses to head compared to torso cooling during explosives ordnance disposal activity in moderate and hot conditions. Thesis, Conventry University, 2011.
2. Buller, M. J., Tharion, W. J., and et al. Estimation of human internal temperature from wearable physiological sensors. In *IAAI* (2010).
3. Epstein, Y., and Moran, D. S. Thermal comfort and the heat stress indices. *Industrial health 44*, 3 (2006), 388–398.
4. Gaura, E., Kemp, J., and et al. Leveraging knowledge from physiological data: On-body heat stress risk prediction with sensor networks. *Biomedical Circuits and Systems, IEEE Trans. 7*, 6 (2013), 861–870.
5. Gaura, E. I., Brusey, J., and et al. Increasing safety of bomb disposal missions: a body sensor network approach. *Trans. Sys. Man Cyber 39*, 6 (2009), 621–636.
6. Kursa, M. B., Rudnicki, W. R., et al. Feature selection with the boruta package, 2010.
7. Niedermann, R., Wyss, E., and et al. Prediction of human core body temperature using non-invasive measurement methods. *Int. J. of biometeorology 58*, 1 (2014), 7–15.

8. Salim, F., Prohasky, D., and et al. Design and evaluation of smart wearable undergarment for monitoring physiological extremes in firefighting. In *Proc. of the 2014 ACM ISWC:Adj*, ACM (2014), 249–254.

9. Talavera, G., and et al. *Protecting Firefighters with Wearable Devices*, vol. 7656 of *LNCS*. Springer Berlin Heidelberg, 2012, book section 65, 470–477.

10. Teunissen, L., Klewer, J., and et. al. Non-invasive continuous core temperature measurement by zero heat flux. *Physiological measurement 32*, 5 (2011), 559.

11. Xu, X., Karis, A. J., and et al. Relationship between core temperature, skin temperature, and heat flux during exercise in heat. *European Journal of Applied Physiology 113*, 9 (2013), 2381–2389.