Proceedings of the 28th IEEE
EMBS Annual International Conference
New York City, USA, Aug 30-Sept 3, 2006

Regularization of Body Core Temperature Prediction during Physical Activity

Andrei Gribok, Thomas McKenna, and Jaques Reifman

Abstract—**This paper deals with the prediction of body core temperature during physical activity in different environmental conditions using first-principles models and data-driven models. We argue that prediction of physiological variables through other correlated physiological variables using data-driven techniques is an ill-posed problem. To make predictions produced by data-driven models accurate and stable they need to be regularized. We demonstrate on data collected during laboratory study that data-driven models, if regularized properly, can outperform first-principles models in terms of accuracy of core temperature predictions. We also show that data-driven models can be made "portable" from one subject to another, thus, making them a valuable, practical tool when data from only one subject is available to "train" the model.**

I. Introduction

Recent advances in physiologic status monitoring resulted from the development of new biosensors and information processing capabilities. These advances have a direct impact on how closely a state of a person can be monitored during civilian activities or during military operations. The new technological capabilities also make possible predictions and estimations of many vital physiological variables, such as body core temperature, heart and respiratory rates and even such subtleties as level of alertness and fatigue. The technological breakthroughs in the development of hardware and firmware are also accompanied by equally profound and significant progress in such fields as data mining and machine learning. Having been equipped with new technology to collect and store relatively large amounts of physiological data in the field, researchers now can explore new horizons in data-driven methods to forecast physiological responses.

For example, the Warfighter Physiological Status Monitoring (WPSM) program seeks to develop a soldier-wearable, computer-based system for providing commanders and medics with critical physiological status information about their war fighters [1], [2]. The WPSM system has two primary aims: the first is to prevent non-battle injuries, such as heatstroke and altitude sickness. A case in point is the approximately 120 heatstroke/sunstroke injuries that occur per year and the associated $10M/yr cost [3]. The second goal is to optimize casualty management through improved casualty detection, diagnostics, and triage. These goals require an array of sensors, a personal area network, and data management software as well as a variety of decision-support algorithms for monitoring and predicting the soldier's physiologic status.

In this paper, we focus on the development and testing of algorithms for non-battlefield injuries and, in particular, on regularization techniques that can be used to enhance the accuracy of prediction of body core temperature.

II. Methods and Techniques

A. SCENARIO Model

The first-principles SCENARIO model [4], a computer-based simulator developed by the U.S. Army Research Institute of Environmental Medicine (USARIEM), predicts and estimates core temperature, heart rate, and other physiological variables. The underlying model for SCENARIO simulates the time course of core temperature, while taking into account different clothing ensembles, workloads, anthropometric characteristics, such as body weight, stature, body fat, fitness, and effects of progressive dehydration. Temperature distribution in the human body is represented by a lump parameter model consisting of six concentric cylindrical compartments. Heat flow is then modeled by a set of macroscopic energy conservation equations, which are based on heat convection between the central blood compartment and the adjacent core, muscle, fat and vascular skin compartments; radial heat conduction between every pair of adjacent compartments; and air convection, radiation and sweat evaporation between the superficial avascular skin layer and the environment and transition through the clothing [4]. These relationships are represented by a set of six ordinary

Manuscript received April 1, 2006. This work was supported, in part, by the Combat Casualty Care and the Military Operational Medicine research programs of the U.S. Army Medical Research and Materiel Command, Fort Detrick, Maryland. In collecting the data presented in this manuscript, the investigators adhered to the policies for protection of human subjects as prescribed in Army Regulation 70-25, and the research was conducted in adherence with the provisions of 45 CFR Part 46. The subjects gave their informed consent for the laboratory study after being informed of the purpose, risks, and benefits of the study.

A. Gribok is with the Bioinformatics Cell, Telemedicine and Advanced Technology Research Center, U.S. Army Medical Research and Materiel Command, Frederick, MD 21702 and also with the Nuclear Engineering Department of the University of Tennessee, Knoxville (e-mail: agribok@bioanalysis.org).

T. McKenna is the Deputy Director of Bioinformatics Cell, Telemedicine and Advanced Technology Research Center, U.S. Army Medical Research and Materiel Command, Fort Detrick, MD 21702 (e-mail: tmckenna@bioanalysis.org).

J. Reifman is a Senior Research Scientist and Director of the Biotechnology HPC Institute, U.S. Army Medical Research and Materiel Command, Fort Detrick, MD 21702 (e-mail: Jaques.reifman@us.army.mil).

differential equations, which can be expressed as follows:

$$\frac{dT}{dt} = A(t) \cdot T(t) + B(t) \qquad (1)$$

where $T(t) \in R^{6x1}$ is a vector representing the temperatures in each of the six modeled compartments of the body, $A(t) \in R^{6x6}$ is a time varying matrix determined by parameters, such as the conductance between two adjacent compartments and blood flow between the compartments. The vector $B(t) \in R^{6x1}$ may be viewed as representing the secondary inputs to the system, and is primarily governed by the metabolic rate in each compartment as well as the respiration rate. Since the data are collected at discrete points in time, in SCENARIO, equation (1) is represented by approximating the temperature gradient by a difference equation.

The need to represent between-subject variability can be addressed by developing data-driven models that utilize historic and real-time data that are specific to the individual. For example, the benefits of a hybrid approach to core temperature modeling have recently been explored where SCENARIO was augmented with a data-driven model [5], [6].

B. Data-driven Models

Data-driven linear models have been used for time series predictions since the early 1970's. The most widely used linear models, such as autoregressive moving average (ARMA) and autoregressive with exogenous inputs (ARX), can be described by the following equation [7]:

$$A(q) \cdot y(t) = B(q) \cdot u(t - nk) + C(q) \cdot e(t) \qquad (2)$$

where A, B and C are polynomials in the delay operator q^{-1}, y(t) and u(t-nk) represent output and input of the system, respectively, at times t and t-nk, and e(t) denotes white noise with finite variance. This generalized equation (2) is known as an ARMAX (which is a combination of ARMA and ARX) model, and produces different autoregressive models for different combinations of A, B and C. For example, if C(q)=1, we obtain an ARX model:

$$A(q) \cdot y(t) = B(q) \cdot u(t - nk) + e(t) \qquad (3)$$

and with no inputs u(t-nk) and with C(q)=1 we obtain an AR model:

$$A(q) \cdot y(t) = e(t) \qquad (4)$$

Other models (ARMA, Moving Average (MA)) can also be obtained by appropriately setting A, B or C to 1.

The models of the ARMAX family predict future values of the target or output variable y as a linear combination of current and delayed inputs, along with current and delayed outputs.

Applying ARMAX to time series predictions involves three steps: selection of model type (AR, MA, ARX, etc.), selection of the order of the polynomials A, B and C, and finally estimation of the coefficients of those polynomials. The model type is normally selected empirically based on prior knowledge of the modeled time series and experiments. The model's order is selected more rigorously by using analytical criteria, such as the Akaike Information Criteria [8], the Minimum Description Length approach [9], or cross-validation. The estimation of ARMAX coefficients is the final step, however, many physiological variables are inherently correlated among each other making this estimation problem an ill-posed one requiring special consideration.

There are different ways in which data-driven models can be applied to physiological data. We consider three approaches arranged in the order of practical usefulness.

1. *Subject Specific* —each subject has an individually tuned model, which is based on tuning the model's type, order and parameters to that individual's data and the model is then applied to that individual only.
2. *Fixed Order*—the model's type and order is fixed beforehand, and only the parameters are tuned to an individual's data and the model is applied to that individual only.
3. *Portable*—the model's type, order and parameters are selected and tuned to one individual's data and the model is applied to other individuals.

From a practical point of view, the third approach is the most valuable one, as it allows the development of models based on one subject's data or based on a limited population of subjects to be applied to a wider population that is not a part of the "training" data set. However, the third approach is also the most challenging one, as it requires developing models with very good generalization capabilities that, in general, constitutes an ill-posed problem.

III. Ill-posed Problems and Physiological Predictions

In 1923 the French mathematician Hadamard introduced the notions of well–posed and ill–posed problems [10]. Hadamard defined a well–posed problem as a problem that satisfies the following three conditions:

a. The solution for the problem exists (existence)
b. This solution is unique (uniqueness)
c. This solution is stable under small perturbation of the data (stability)

If any one of these conditions is not met, the problem is said to be ill-posed and requires special consideration.

In physiological prediction, one or more of these conditions may be violated. First, the variables related to core temperature may not provide all of the information that is necessary to estimate a true value of the core temperature. Second, different variables can provide different information about the value of the core temperature, and, finally, core temperature prediction can be unstable due to small perturbations in the input data. For example, using slightly different number of training data samples may yield significantly different predictions. In the current study, we assume that the variables selected to infer the value of the core temperature do provide information about this value. We also assume that this information is full, and unambiguous, thus, postulating the existence and uniqueness of the solution for the core temperature prediction. Our primary concern is to maximize the stability or consistency of the core temperature prediction, and we shall show that stability and consistency can be assured by using regularization.

A very popular method to estimate coefficients of data-driven models, including ARMAX models, is the methods of ordinary least squares, which can be written in its generic form as

$$\arg\min_{w} \|Xw - y\|_2^2, \quad X \in R^{mxn}, \ m \geq n \tag{5}$$

where X is the design matrix of predictor variables, y is the response variable (core temperature in our case) and w is the sought regression coefficients. As mentioned before, the physiological variables are highly correlated among each other, thus, producing ill-conditioned or even rank deficient design matrices.

To deal with ill-conditioned problems having ill-determined numerical rank, the method of regularization proposed by Tikhonov [11] can be used. In this method, the minimization problem (5) is replaced by the following augmented functional:

$$\arg\min_{w} (\|Xw - y\|_2^2 + \lambda^2 \|Lw\|_2^2) \tag{6}$$

The regularization parameter λ controls the trade-off between the smoothness of the solution and its fit to the data, and L is a well-conditioned matrix; for example, a discrete approximation of a derivative operator. The regularization parameter λ, in this study, was selected using generalized cross-validation [12].

IV. Results and Discussion

To show how regularization affects predictions of physiological values we applied unregularized and regularized ARX models to core temperature measurements obtained in laboratory experiments. The laboratory data set was based on nine volunteer subjects [13] (Age:23±4 yr; Height:174.2±5.8 cm;Weight:73.4±6.5 kg; Body Fat Pct: 17.9±3.99 %). The subjects walked on a treadmill under two environmental conditions: (i) CONTROL (Day 1) (20°C temperature and 50% relative humidity); and (ii) HUMID (Day 2) (27°C temperature and 75% relative humidity). The wind speed was 1.1 m/s (2.5 mph) for both conditions. On the morning of test days, the subjects, dressed in air permeable battle dress uniform, were instrumented for the collection of various physiological variables, including core (rectal) temperature. Then they sat on a chair for 10 minutes just before starting to walk at 3 mph on level treadmills. The walking paused after every 30 minutes for 10 minutes of sitting. There were four 30-minute walking periods per test, so that the entire experiment lasted a total of 170 minutes, including 10-minute rest periods at each end. At the end of each 10-minute pause, the subjects were given 150 ml of water before walking again. Rectal temperature (assumed to be representative of the core temperature) was collected continuously and recorded every minute.

Seventeen variables, corresponding to anthropomorphic and environmental conditions, were used as inputs to SCENARIO. Based on these data, SCENARIO produced estimates of the core temperature during the study. An example of SCENARIO prediction of the core temperature of subject 1 along with the activity profile and actual core temperature measurements is shown in Fig. 1.

To implement the data-driven model, seven variables (age, height, weight, body fat percentage, mean radiant temperature, relative humidity and walking speed) were selected, from the original set of 17, and used as inputs, along with the delayed output of the core temperature measures.

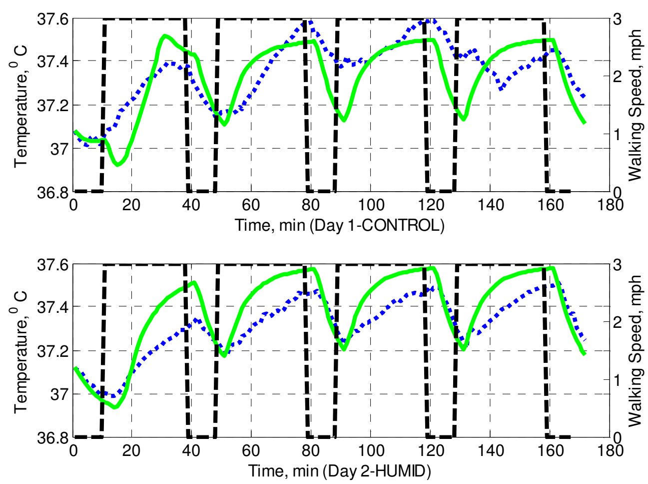

Fig.1 Simulations of the core temperature of subject 1 using SCENARIO model (solid) against actual core temperature measurements (dotted) and activity profile (dashed).

After extensive trials of different model types (ARMA, AR, ARX), the ARX model was selected as the one that describes the data most accurately. The order of the polynomials in the ARX model was selected using cross-validation, i.e., the first half of the data from the CONTROL environmental conditions together with the first half of the data from the HUMID environmental conditions were used as training data to estimate parameters of different order ARX models. The models were then used to make core temperature predictions over the last half of the data (i.e., the testing data) for each of the environmental conditions. The model that produced the smallest validation error over the testing data was selected as an optimal model for that subject. It should be stressed that while the SCENARIO model estimates core temperature over the entire range from one prediction time step to the next, the data-driven model predicts core temperature for a given time horizon. In the simulations shown here we set the prediction time horizon at 20 minutes, as this represents a reasonable and useful time interval for an individual to adjust his activities to avoid heatstroke.

Figure 2 shows core temperature predictions obtained with an unregularized ARX model. The models' performance was evaluated using root mean squared error (RMSE) estimated only for the test portion of the data. Notice the highly oscillatory nature of the ARX predictions, which is a manifestation of the instability of the ARX model. Also, notice that for Day 2 the data-driven model is only marginally better than SCENARIO in terms of RMSE. Figure 3 shows the core temperature prediction using the same data and a regularized ARX model. The change is dramatic, with the ARX model now producing consistent and smooth predictions. The RMSE for the regularized ARX model is significantly smaller than that of the unregularized model and several times smaller than SCENARIO's RMSE computed over the same test portion of the data. It is worth noting that, when compared with the unregularized model, the regularized model produced an inferior fit for the training portion of the data, especially for Day 2. However, this was compensated with superior fit on the test portion of the data. This represents a typical trade-off for regularized models, where the training error is increased at the expense of a reduced testing error and improved generalization.

Data-driven modeling is most useful if data from one subject can be used to train a model and this very same model can be used to reliably predict the responses of other subjects. To check whether this holds for regularized models, we trained an ARX model on one subject's data and applied this model to predict other subjects' data. The results of such predictions, with a 20-minute prediction horizon, are presented in Fig. 4.

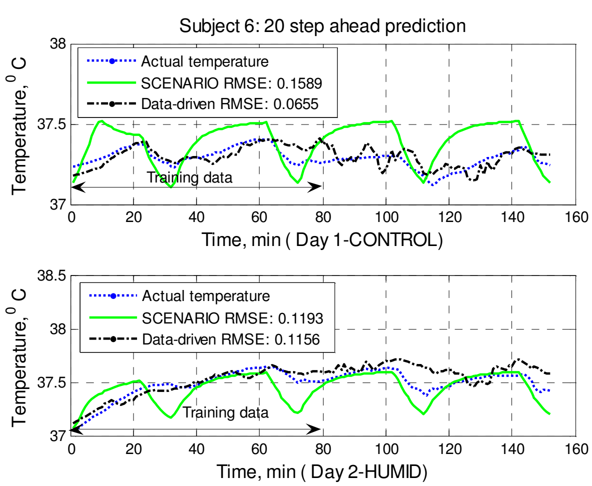

Fig.2 Core temperature prediction for subject 6 using an unregularized ARX model.

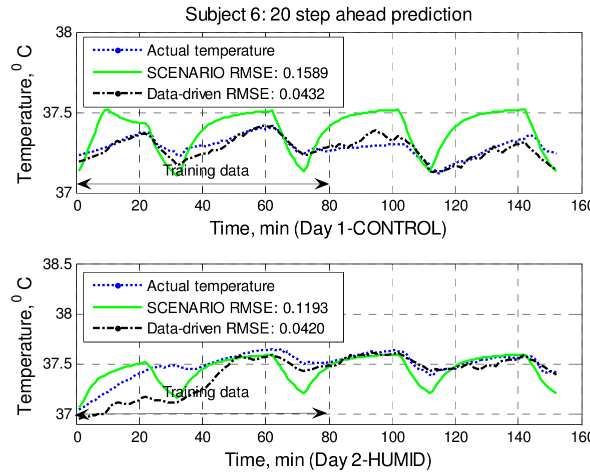

Fig.3 Core temperature prediction for subject 6 using a regularized ARX model.

The bar columns in Fig. 4 correspond to RMSEs obtained for SCENARIO and the three different modeling approaches discussed in Section II. Results are provided for each of the nine (1-9) subjects plus an overall average over all subjects, shown in the last column. The SCENARIO result for each subject indicates averages of SCENARIO estimations for the two days over eight subjects, with that subject's estimation excluded from the averaging. Each SCENARIO simulation for a given subject, however, was performed with anthropometric data for that subject. The results shown for the data-driven Portable models were similarly computed. A model was optimized and trained on one subject's data and subsequently applied to all other subjects. So, the result for each subject indicates the average RMSE over the other eight subjects, excluding the one for which the model was developed for. The data-driven Fixed Order model corresponds to approach 2 described in Section II, wherein the order of the data-driven model is fixed while the coefficients are adjusted to each individual and the model is tested on that individual's data. Finally, the Subject Specific results correspond to approach 1 in which the order of the data-driven models, along with its parameters, are tuned and tested using the data from the same individual. For the case of the Fixed Order and Subject Specific data-driven models, the RMSEs represent the sum of testing RMSEs obtained on the second halves of both days while the models were trained on the first halves. All data-driven models were properly regularized.

Not surprisingly, the smallest RMSEs resulted from Subject Specific models, where data from the same subject were used for training and testing and the model's order and parameters were tuned to that individual. The second smallest error is produced by the data-driven model that used a fixed order for all the subjects but the parameters were adjusted to specific individuals.

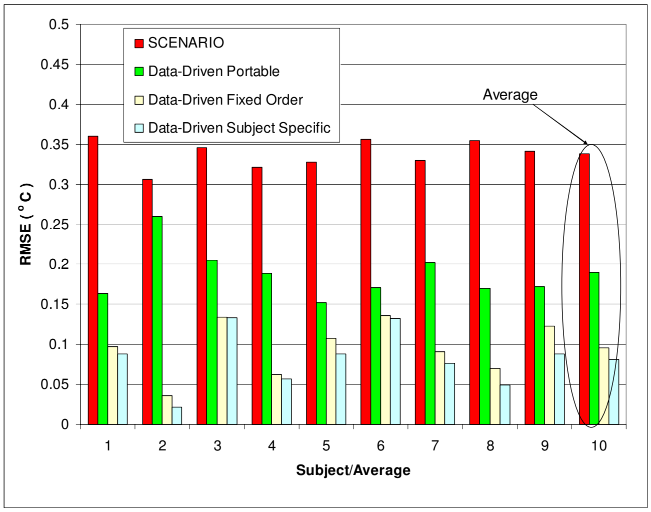

Fig.4 Comparison of predictive performance of different modeling techniques.

The most important finding, however, is that RMSEs for the Portable models are smaller than SCENARIO's RMSEs. This fact shows that data-driven models can be effectively applied when data from only one subject are available for training. This is a very valuable result for practical applications, as it may obviate the need to customize models for each subject, while facilitating the incorporation of individualized models as part of the WPSM system. It should be stressed, however, that there is a fundamental difference in which the data are used in the two approaches. The data-driven models use the currently available measurements to make further predictions while SCENARIO does not make use of such information.

V. CONCLUSIONS AND FUTURE WORK

Data-driven prediction of physiological variables through the use of other correlated physiological variables may result in an ill-posed problem. Regularization alleviates this problem and assures that data-driven models provide stable and consistent predictions and can even be "portable" from one subject to the other. Properly regularized data-driven models can, on average, outperform first-principles models across the entire testing population, as illustrated in this initial study employing a controlled laboratory data set.

These results, however, are preliminary and based on a cyclical laboratory exercise regime that is not representative of typical field operations. To verify our key finding, that data-driven models can be made "portable" from one individual to another, we have initiated more rigorous testing employing field-collected data representative of actual soldier activities.

ACNOWLEDGMENT

The authors would like to acknowledge the support of the following people: N. O. Oleng' for help with SCENARIO simulations; W. Santee and L. Berglund for the laboratory study data. The laboratory study was funded, in part, by the Natick Soldier System Center. We also would like to thank M. Kolka and L. Berglund for their useful comments and suggestions on the manuscript.

DISCLAIMER

The opinions or assertions contained herein are the private views of the authors and are not to be construed as official or as reflecting the views of the U.S. Army or of the U.S. Department of Defense.

REFERENCES

[1] R.W. Hoyt, J. Reifman, T. S. Coster, and M. J. Buller, "Combat medical informatics: present and future," in Proc. *AMIA 2002 Annual Symposium*, San Antonio, Texas, 2002, pp 335-339.

[2] R.W. Hoyt, and K. E Friedl, "Current status of field applications of physiological monitoring for the dismounted soldier," in *Metabolic Monitoring Technologies for Military Field Applications.* M. Poos (Ed.). National Academy of Sciences, National Academy Press, 2004, pp. 247-257.

[3] M. N. Sawka, C. B. Wenger, and K. B. Pandolf, "Thermoregulatory responses to acute exercise-heat stress acclimation," in *Handbook of Physiology,* Section 4: Environmental Physiology. C.M. Blatteis and M. J. Fregley, Eds., New York: Oxford University Press for the American Physiological Society. Ch. 9, pp. 157-185,1996.

[4] K. K. Kraning and R. R. Gonzalez, "A mechanistic computer simulation of human work in heat that accounts for physical and physiological effects of clothing, aerobic fitness, and progressive dehydration," *J. Therm. Biol.,* vol. 22, pp. 331-342, 1997.

[5] N. O. Oleng', J. Reifman, L. Berglund, and R.W. Hoyt, "Alternative approaches to improve physiological predictions," in *Proceedings of the Army Science Conference,* Orlando, FL, KO-05, 2004.

[6] N. O. Oleng', and J. Reifman, "Hybrid Approaches to Physiologic Modeling and Prediction. Conference on Bio-monitoring for Physiological and Cognitive Performance during Military Operations," in *Proc. of the SPIE Symposium on Defense and Security,* Orlando, FL, 2005, vol. 5797, pp. 93-203.

[7] L. Ljung, *System Identification – Theory for the User*. NJ: Prentice Hall, 1999.

[8] Y. Sakamoto, M. Ishiguro, and G. Kitagawa, *Akaike Information Criterion Statistics*. D. Reidel Publishing Company, 1986.

[9] J. Rissanen, "Modeling by shortest data description," *Automatica,* vol. 14, pp. 465-471, 1978.

[10] J. Hadamard, *Lectures on Cauchy's Problem in Linear Partial Differential Equations,* Yale University Press, New Haven, 1923.

[11] A. N. Tikhonov, "Solution of incorrectly formulated problems and the regularization method," *Doklady Akad. Nauk USSR,* 151, pp. 501-504, 1963.

[12] G. Golub, M. Heath, and G. Wahba, "Generalized cross-validation as a method for choosing a good ridge parameter," *Technometrics*, vol. 21, pp. 215–223, 1979.

[13] W. R. Santee, L. G. Berglund, A. Cardello, C. A. Winterhalter, and T.L. Endrusick, "Physiological and Psychological Assessment of Volunteers Wearing Air Permeable Battle-Dress (BDU) Uniforms during Intermittent Exercise, U.S. Army Research Institute of Environmental Medicine," Technical Report (in preparation).