Hybrid Approaches to Physiologic Modeling and Prediction

Nicholas O. Oleng' and Jaques Reifman

Bioinformatics Cell
Telemedicine and Advanced Technology Research Center
U.S. Army Medical Research and Materiel Command
Frederick, MD 21701

ABSTRACT

This paper explores how the accuracy of a first-principles physiological model can be enhanced by integrating data-driven, "black-box" models with the original model to form a "hybrid" model system. Both linear (autoregressive) and nonlinear (neural network) data-driven techniques are separately combined with a first-principles model to predict human body core temperature. Rectal core temperature data from nine volunteers, subject to four 30/10-minute cycles of moderate exercise/rest regimen in both CONTROL and HUMID environmental conditions, are used to develop and test the approach. The results show significant improvements in prediction accuracy, with average improvements of up to 30% for prediction horizons of 20 minutes. The models developed from one subject's data are also used in the prediction of another subject's core temperature. Initial results for this approach for a 20-minute horizon show no significant improvement over the first-principles model by itself.

Keywords: hybrid modeling, black-box model, neural networks, physiological modeling

1. INTRODUCTION

In recent years, there have been tremendous advancements in bio-monitoring technology, particularly in military medicine applications. New bio-sensors and greater information processing capability now permit on-line, real-time measurement of physiological variables. In order to fully utilize these capabilities in modeling and predicting physiological variables, it is necessary to investigate data-driven algorithms that can potentially provide greater fidelity than the ones currently in use. This stems from the fact that data-driven models, taking physiologic data from the very same subject we wish to develop a model for, would directly capture a subject's physiologic variability; which is a phenomenon that has eluded existing conventional modeling approaches.

The vulnerability of warfighters to heat injuries during periods of extreme temperatures is a significant problem in the military. Of particular concern is heat stroke. For example, in 2002, there were 1816 heat-related injuries of active duty soldiers in the US Army.[1] This number could be significantly reduced if we were able to rapidly identify vulnerable individuals, so that appropriate action could be taken to prevent such injuries. There is general consensus that in order to monitor this preventable injury, accurate measures and predictions of core temperature are required.

In this paper, we are primarily concerned with the prediction of core temperature. Specifically, we demonstrate how a hybrid technique, which combines a first-principles model with a data-driven model, can be used to improve the prediction of core temperature of individual subjects. This approach is tested using laboratory data with a view to eventual application on field data. The technique, if successful, could be employed in models for predicting heat strain, and incorporated into the Warfighter Physiological Status Monitoring (WPSM) project.[2,3] The ultimate goal of the WPSM project is to develop a suite of soldier-wearable sensors and decision support algorithms to provide commanders and medics in the field with critical physiologic status information about their warfighters. In particular, the ability to prevent non-battle injuries, such as heat stroke, is of key importance.

Physiological models commonly rely on first-principles knowledge about various mechanisms in the human body and their associated dynamics. The resulting models may be effective in capturing average physiological

Further author information: (Send correspondence to Nicholas Oleng'.)
Nicholas Oleng': E-mail: nicholas@bioanalysis.org
Jaques Reifman: E-mail: jaques.reifman@amedd.army.mil

responses, but are not necessarily effective in modeling a particular individual, leading to inaccurate predictions for that individual. The capacity to "tune" a model to a specific individual is particularly important due to the well-documented inter-subject variability.[4,5] Individuals with similar anthropomorphic characteristics, subjected to the same workload and environmental conditions, may yield very different physiological responses. Such variation in physiologic response is especially critical at limiting thresholds of physiologic health,[6,7] such as extreme values of core temperature, where small variations (as little as 0.5^0 C) can make the difference between a suitable recovery and an irreversible pathological condition. The need to represent this variability can be addressed by developing models that utilize historic and real-time data that are specific to that individual.

SCENARIO,[8] a predictive thermal model developed by the U.S. Army Research Institute of Environmental Medicine (USARIEM) was designed to predict core temperature, heart rate and other physiological variables from environmental, clothing, anthropometric and physical activity inputs. The SCENARIO model relies primarily on knowledge provided by fundamental first principles of physics and physiology. While the model does incorporate parameters, such as an individual's weight, height and fat percentage, its fidelity could be improved by accounting for additional relevant processes of human physiology. One way of improving model fidelity, is to incorporate data-driven or "black-box" models into the first-principles model, creating "hybrid" models. Black-box models have proven to be effective in other applications.[9,10] Examples of black-box models that could be used in prediction include autoregressive with exogenous input (ARX) models and neural network models.

In this paper, we augment the prediction provided by the SCENARIO model with the prediction from each of the two black-box models mentioned above, applied separately. Such additional predictions are expected to account for physiological processes as well as individual variability not captured by the SCENARIO model.

2. THE SCENARIO MODEL

The underlying model for SCENARIO simulates the time course of core temperature, while considering 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, considering heat convection between the central blood compartment and the adjacent core, muscle, fat and vascular skin compartments; radial heat convection between every pair of adjacent compartments; and air convection, radiation and sweat evaporation between the superficial avascular skin layer and the environment, through the clothing.[8] These are represented by a set of six ordinary differential equations which can be expressed as follows:

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

where $T(t) \in R^{6\times1}$ is a vector representing the bulk temperatures in each of the modeled compartments of the body, $A(t) \in R^{6\times6}$ is a time varying matrix determined by parameters, such as the conductance between two adjacent compartments and blood flow between the compartments. The vector $B(t) \in R^{6\times1}$ may be viewed as the secondary inputs to the system, and is governed by the metabolic rates in each of the compartments 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 inputs needed for running the SCENARIO model include:

1. environmental: mean radiant temperature, ambient temperature, vapor pressure, wind speed
2. activity based: walking speed, pack weight (load), terrain factor, slope/grade
3. anthropometric: age, weight, height, fat percentage, water intake
4. clothing insulation and permeability.

It is evident that for estimates of metabolic rates to be of any significance in accurately predicting core temperature, they must be based on parameters specific to an individual, such as age, body weight and activity level. In this respect, SCENARIO provides more accurate predictions than would group-average models. However, group average values are used to estimate some of parameters of the model, such as tissue conductance and heat capacity, since it would be difficult to do so for every subject. Additionally, since all parameters are estimated on the basis of experimental data, inherent observation error and limited sample size may lead to discrepancies that, compounded, could contribute to model inaccuracy. Furthermore, because of the simplifying characteristic of the SCENARIO model, some of the physiological dynamics may not have been captured.

3. THE HYBRID APPROACH TO MODELING

The traditional approach to physiological modeling has been the development of parametric models derived from prior knowledge, in the form of empirical correlations, known mathematical equations or fundamental first principles.[11] In order to produce accurate predictions, this approach requires considerable prior knowledge about the underlying phenomena being modeled. However, as with any natural process, a complete understanding of physiological phenomena is not attainable. As a result, these models often fail to make accurate predictions across the entire spectrum of operation.

On the other hand, nonparametric models have been used to model complex processes when exact analytical equations are unavailable or difficult to develop.[10] These models are developed from process data, where the functional form of the model is conformed to the specifics of the particular process only after the presentation of the data. There are a wide variety of such models ranging from simple ARX models to more complex and nonlinear neural network models. These data-driven models are nevertheless limited to making predictions within the boundaries of the training data. Their usefulness is extremely dependent on the quantity and quality of the training data, which are often difficult to obtain and verify.

More recently, hybrid approaches that combine first-principles models and the data-driven models mentioned above have been proposed as alternative, more flexible, and perhaps superior modeling paradigms than more traditional parametric and nonparametric approaches.[12,13] The promise of hybrid approaches lies in their potential to use the best of both approaches while avoiding the limitations of each approach used separately. Data-driven models complement missing first-principles knowledge with information extracted directly from the process data, while first-principles-based models compartmentalize and reduce the role of data-driven models to specific functionalities, significantly reducing the training data requirements and improving model generalization and extrapolation.

In the following subsections, we provide brief descriptions of both ARX and neural networks models. In addition, we describe how these models are used in conjunction with SCENARIO, to provide better predictions of core temperature.

3.1. Linear (ARX) Models

The ARX model used in our investigations is of the general form:[9]

$$y(t+d) = \phi^T(t)\,\theta + n(t), \tag{2}$$

where the regression vector $\phi(t)$ consisting of the past n_a outputs and the past n_b inputs is defined as

$$\phi^T(t) = \begin{bmatrix} y(t) & y(t-1) & y(t-2) & \cdots & y(t-n_a) & u(t+d-1) & u(t+d-2) & \cdots & u(t-n_b) \end{bmatrix}. \tag{3}$$

Here, $y(t)$ represents the output of the model at time t, d corresponds to the delay of the system (typically 1 in our case), $u(t)$ represents the input at time t, and $n(t)$ denotes a white Gaussian noise signal. In addition, θ is the vector of parameter coefficients associated with the regression vector $\phi(t)$.

In practice, the estimate $\hat{\theta}$ of the parameter vector θ is obtained by fitting the model (2) to training data $\{y_1(t), u_1(t)\}$, i.e., a record of past inputs $y_1(t)$ and outputs $u_1(t)$. The most common form of parameter estimation is the least squares optimization technique in which $\hat{\theta}$ is computed so as to minimize a performance measure given by

$$J = \sum_t e_1(t)^2, \tag{4}$$

where $e_1(t) = y_1(t) - \hat{y}_1(t)$ and $\hat{y}_1(t)$ is the estimate of the output $y(t)$ provided by the ARX model. To test the validity of the model, a new set of data $\{y_2(t), u_2(t)\}$ previously unused in the training is required. The resulting errors $e_2(t) = y_2(t) - \hat{y}_2(t)$, where the estimates $\hat{y}_2(t)$ are obtained by applying $u_2(t)$ to the model with $\hat{\theta}$, provide a measurement of the goodness of fit of the model.

In implementing the model for prediction, the actual measurements of past outputs and inputs are used in the regression vector $\phi(t)$ and a prediction $\hat{y}(t+d)$ of $y(t+d)$ obtained. To predict the value of $y(t+d+H)$, where $H > 0$, the prediction is performed iteratively H times, so that $\hat{y}(\tau) : \tau \in [t+d+H+1, t+d+H-1]$ is used in the corresponding regression vector instead of the unavailable $y(\tau)$. For further information on the use of such linear models, the reader is referred to the books by Ljung[9] and Goodwin and Sin.[10]

3.2. Neural Network Models

The model described in the previous subsection is useful when the underlying process is linear. However, when the process becomes nonlinear, such models are no longer able to provide accurate representations of the process and a more complex model is required. Neural networks have proven useful in modeling processes with nonlinear dynamics. A neural network model for such a system can ostensibly be viewed as a nonlinear version of the ARX model in Equations (2) and (3), so that

$$y(t+d) = f(\phi(t)) + n(t), \tag{5}$$

where $\phi(t)$ and $n(t)$ are defined as in Equations (2) and (3), and $f(\cdot)$ is the scalar nonlinear function representing the neural network.

There are two main types of neural networks commonly used in modeling applications: feed-forward neural networks, also commonly referred to as multi-layer neural networks, which are static in nature; and recurrent neural networks, which incorporate delays in their architecture and are therefore dynamic in nature. Recurrent networks, therefore, form a more natural approach to modeling dynamic processes and are used in our investigation. Figure 1 provides a pictorial representation of a simple recurrent neural network. Delay in the outputs is captured implicitly in the network by introducing the tap delay line (indicated as z^{-1} in the figure), that ensures that past values of the outputs from the outermost layer are fed back into the input layer. The result is that the output of the network is a function of past outputs and inputs, even when these have not been explicitly fed into the network through the input layer. Feed-forward neural networks can, in principal, also model dynamic processes, but they require the explicit incorporation of past inputs and past outputs, as inputs to the network. Recurrent networks on the other hand, simply take in present inputs but can store these as well as past outputs within their architecture. The neural networks are commonly trained using the back-propagation algorithm,[14] which minimizes a performance criterion similar to the one in Equation (4).

In this paper, we employ Elman recurrent neural networks and variations of the back-propagation training algorithm (available in the MATLAB Neural Network toolbox[15]).

3.3. Implementing the Hybrid Approach for SCENARIO

The hybrid approach, when implemented in the context of core temperature prediction involving the SCENARIO model, allows us to employ prior knowledge about the "human physiological process" to the maximum extent possible, and complements the missing knowledge with information extracted from core temperature measurements.* More specifically, the black-box model employs online feedback of subject-specific data into the hybrid model, ensuring improvements in predictions.

In this investigation, we employ the parallel approach to hybrid modeling, whereby the data-driven, black-box model is aimed at estimating the deviation of SCENARIO predictions from the actual temperature measurements. The estimation mechanism is enhanced by the feedback of the prediction errors into the black-box model as shown in Figure 2. Both black-box models, the ARX and the neural network, are separately employed to estimate the offset between the SCENARIO predictions and the actual core temperature measurements of specific subjects. Accordingly, the resulting prediction of core body temperature is the sum of the predictions provided by the SCENARIO model and that provided by the black-box model represented by either an ARX model or a neural network model.

*For future field applications, the real-time physiologic measurements could be provided by an array of WPSM biosensors.

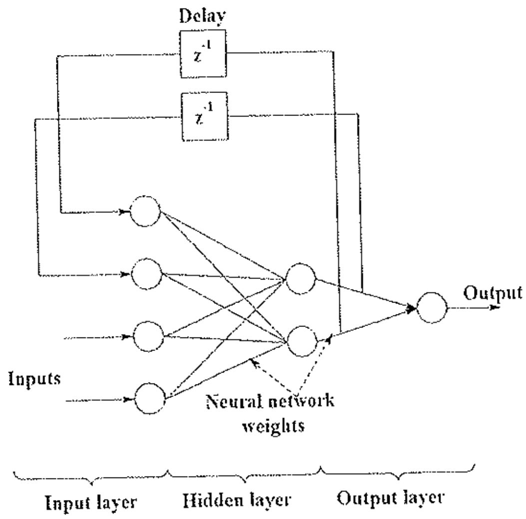

Figure 1. A pictorial representation of a recurrent neural network

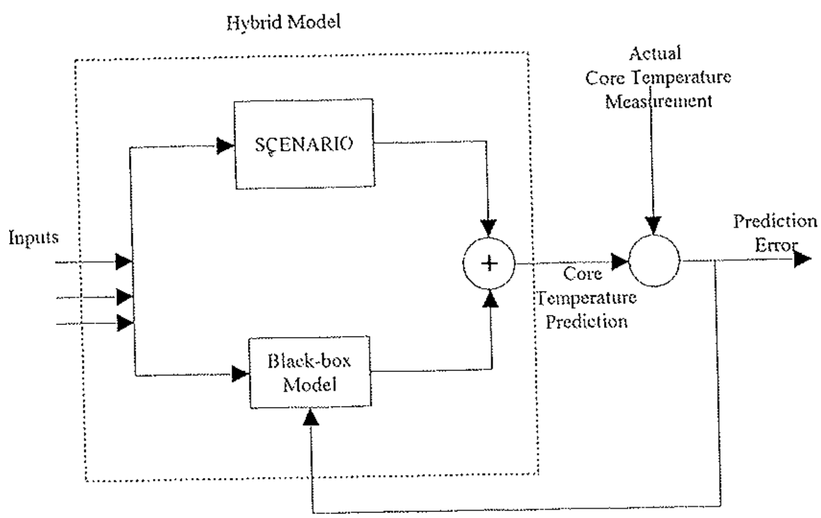

Figure 2. The prediction mechanism

4. VALIDATION AND RESULTS

In this section, we present the results of the implementation of variants of the hybrid model approach to laboratory data using two data-driven models. The particular models implemented here involve: (i) the SCENARIO model in parallel with an ARX model, and (ii) the SCENARIO model in parallel with a neural network model. These models are developed and tested using data obtained from the laboratory study described below.

The Laboratory Study:

Core temperature data from nine volunteers during treadmill walking in two environmental conditions were used in the simulations presented in this paper†. These conditions were: (i) CONTROL (20^0C/68^0F temperature and 50% relative humidity); and (ii) HUMID (27^0C/81^0F temperature and 75% relative humidity). The wind speed was 1.1 m/s (2.5 mph) for both environments. On the morning of test days, the volunteers, dressed in air permeable battle dress uniform (BDU), were instrumented for the collection of various physiological measurements, including core (rectal) temperature. Then they sat on a chair for 10 minutes just before starting to walk 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 was collected continuously and recorded every minute. Heart rate was monitored intermittently for volunteer safety.

The Simulation and Results:

For each of the subjects, 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 in the initial training of the black-box models to obtain the coefficients in the parameter vector θ in Equation (2) as well as the weights for the neural network in Equation (5). The resulting models were then used to make core temperature predictions over the last half of the data for each of the two environmental conditions. A measure of goodness of the predictions was computed as

$$fit = 100\left(1 + 0.5\frac{||\hat{y} - y||}{||y - \bar{y}||}\right)^{-1}, \tag{6}$$

where $\hat{y}$ is the vector of the predicted values, y is the vector of the actual temperature values and $\bar{y}$ is the arithmetic mean of the vector y. Equation (6) is a variation of the measure

$$fit = 100\left(1 - \frac{||\hat{y} - y||}{||y - \bar{y}||}\right), \tag{7}$$

provided by Ljung.[9] The measurement of fit is chosen so that all the measures lie between 0 and 100, where 0 represents the worst possible fit and 100 thé best. This measure was chosen instead of the mean square error (MSE)[17] so as to give a quick insight into the relative ranking of a given prediction. In addition, owing to the small variations in temperature, an anomaly may occur whereby a relatively flat line may provide a small MSE indicating what may seem to be a good prediction of an otherwise varying temperature profile.

The ARX model used in this investigation was given by Equations (2) and (3), with $n_a = 4, n_b = 4$ and $d = 1$. The neural network was chosen to have two layers, the hidden layer having 14 sigmoidal activation functions and the output layer having a single linear activation function. In choosing the size of the models, a section of the data is initially used to develop (train) models of different sizes. As expected, as the size of the models increases the accuracy of the models in predicting the training data increases. When the models are used for predicting new data, this is again observed initially. However, as the size of the models increases beyond a certain limit, an increase in accuracy is no longer observed. This limit is then chosen as an appropriate size for the models since it provides the best tradeoff between size and accuracy.

Figure 3 shows the performance of the hybrid ARX model in relation to the actual core temperature and the prediction by SCENARIO, for subject 1. The dashed (green) line shows the predictions provided by the original SCENARIO model while the (red) dashed line with a cross shows the hybrid model predictions. The actual temperature is shown by the blue (solid) line. These hybrid model predictions are obtained with a time horizon set to 20 minutes. In other words, for each minute, the actual prediction for that minute was computed 20 minutes earlier. The first graph in the figure shows the predictions under the CONTROL condition, while the second shows the predictions under the HUMID condition. While the graphs show the predictions over the entire day, the measures of fit were computed only for the last half of the data for each environmental condition.

†The study[16] was conducted by USARIEM under the principal investigator, Dr. William Santee.

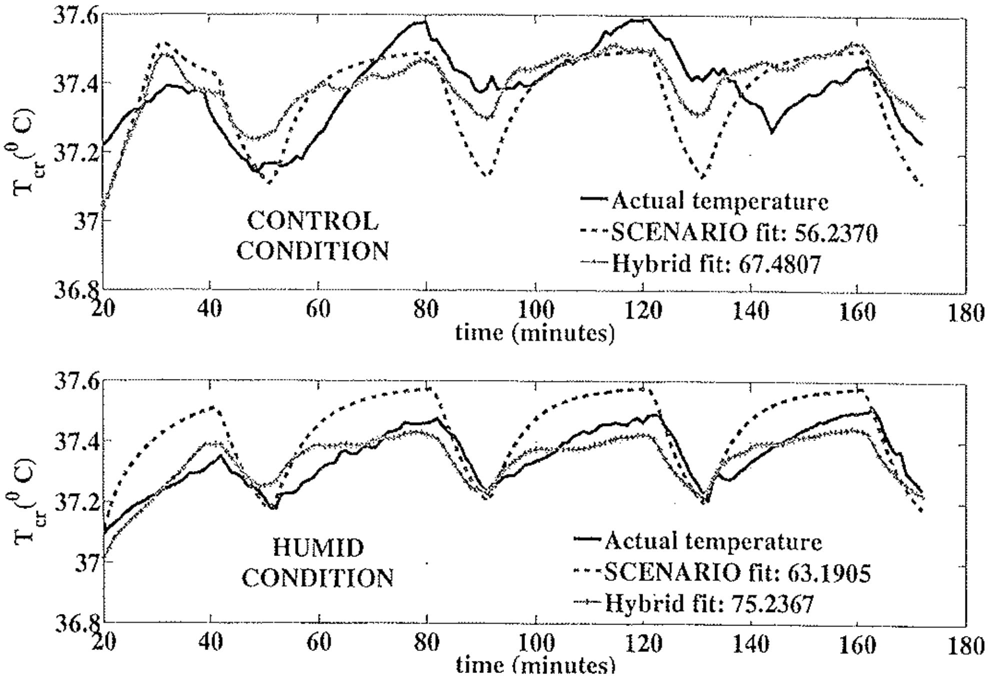

Figure 3. Subject 1: 20-minute ahead prediction of core temperature using an auxiliary ARX model in parallel with SCENARIO

As can be seen from Figure 3, SCENARIO tends to underpredict the core temperature during the rest periods in the CONTROL condition, and to overpredict it during exercise in the HUMID condition. The hybrid models therefore serve to improve prediction by making up for this offset. In this case, the average performance as measured by the fit in Equation (6) is roughly 20% better than that of SCENARIO alone.

Tables 1 and 2 show the performance of the hybrid models in relation to SCENARIO under both CONTROL and HUMID conditions, respectively. The second column in each table shows the performance of the ARX model in parallel with SCENARIO for each subject. Corresponding to this, in the third column, is the performance of the SCENARIO model alone. The mean performance over the nine subjects for both the hybrid approach and SCENARIO are shown in the last row of the tables. As can be seen from the tables, the addition of the black-box model to SCENARIO helps improve the accuracy of the 20-minute-ahead prediction by about 30%.

As the prediction horizon is increased, however, the predictions provided by the hybrid model begin to deteriorate. In addition, there may be an increased phase shift in the hybrid model predictions. This is attributed to the black-box part of the hybrid model. From the description of the prediction technique for data-driven models given in Section 3, it can be seen that an increased prediction horizon leads to a greater accumulation of prediction error, hence the overall deterioration of the model as a function of the prediction horizon. In contrast, the performance of the SCENARIO model by itself, which does not incorporate explicit feedback of the measured core temperature, remains the same regardless of the prediction horizon. This underscores the importance of the first-principles part of the model (SCENARIO in this case) as the time horizon increases. Indeed for large time horizons, the black-box part of the model deteriorates the performance of the hybrid model to the point where the hybrid model accuracy becomes worse than the use of SCENARIO alone.

Figure 4 shows the results for a 20-minute-ahead prediction where a neural network is used as the black-box part of the hybrid model. As can be seen for the figure, there is a jaggedness in the predictions. A possible explanation for this may be that there is not enough data to sufficiently train the neural network. The same phenomenon of jaggedness is observed in predictions for other subjects. However, the accuracy of the neural-network-based hybrid model is still, on average, 15% better than that of SCENARIO alone, for a 20 minute

Table 1. Goodness of fit measures for CONTROL environmental conditions: 20-minute-ahead prediction

Subject ID (i)	Fit when subject i's ARX model is used	Fit of SCENARIO	Average Fit for subject i $\frac{1}{8}\sum_{j=1,j\neq i}^{9} Fit_j$	Fit when subject i's neural network model is used
1	67.4807	56.2370	42.2767	52.0914
2	44.6839	19.7705	28.4385	38.9054
3	44.8539	57.5808	45.2530	56.4232
4	47.9348	22.6074	36.5133	47.3933
5	65.3632	31.0991	27.8910	54.0983
6	28.1898	37.3540	32.6061	34.4323
7	42.2732	27.9527	44.9422	42.0192
8	52.3356	44.5337	36.1877	40.0343
9	42.0285	43.8825	38.4949	43.0433
Mean	48.3493	37.8909	36.9559	45.3823

Table 2. Goodness of fit measures for HUMID environmental conditions: 20-minute-ahead prediction

Subject ID (i)	Fit when subject i's ARX model is used	Fit of SCENARIO	Average Fit for subject i $\frac{1}{8}\sum_{j=1,j\neq i}^{9} Fit_j$	Fit when subject i's neural network model is used
1	75.2367	63.1905	69.3169	73.371
2	45.0048	25.0270	29.4941	43.0534
3	48.6363	36.5026	35.3413	32.0423
4	57.4675	28.8949	40.8591	42.9423
5	59.5394	51.8208	48.4882	42.5232
6	68.1682	53.7107	48.0697	55.0324
7	54.5377	32.5668	44.4374	43.0978
8	39.9021	49.8594	46.1365	34.0232
9	50.0318	36.6813	26.5301	51.0423
Mean	55.3916	42.0282	43.1859	46.3475

prediction horizon. The last column of Tables 1 and 2 shows that the performance of the resulting hybrid model is on average worse than that of the ARX hybrid model. It is expected that with more data, the predictions with the neural network hybrid models would improve and possibly provide better predictions than that of the ARX hybrid model.

The availability of sufficient data for training is key to the development of hybrid models, not only for neural-

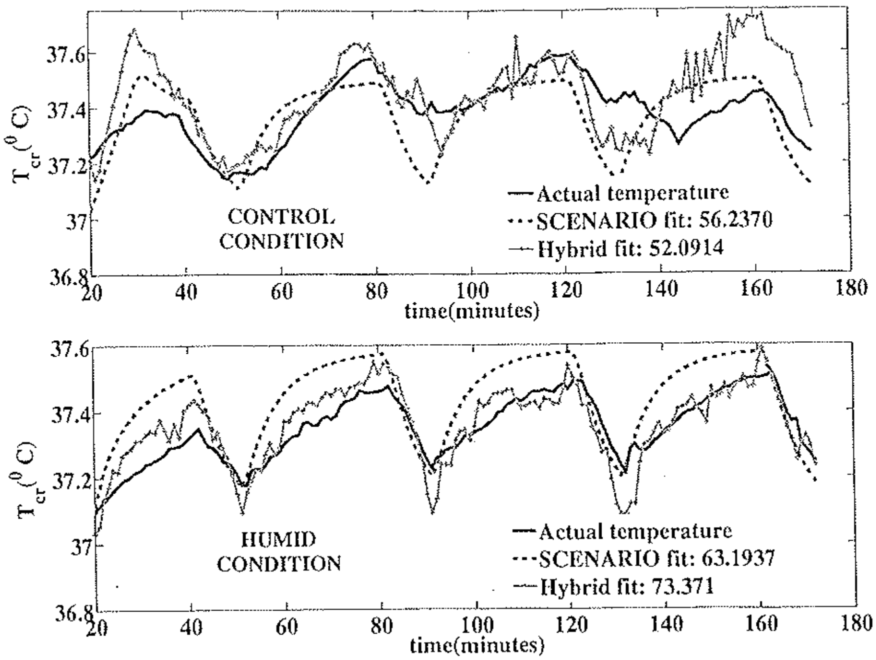

Figure 4. Subject 1: 20-minute ahead prediction of core temperature using an auxiliary neural network model in parallel with SCENARIO

network-based models, but also for ARX-based models. The results shown in this paper are therefore limited by this fact. While we report predictions carried out when the first half the data are used for training and the second half are used for prediction, we have conducted similar investigations in which three quarters (instead of half) of the data were used for training of the ARX-based models. These investigations showed an improvement in performance of up to 20% over the case in which only half the data were used for training. Hence, there is a great need for more data in order to fully utilize the potential of these hybrid models.

Robustness of Models

The issue of robustness of the models developed in this investigation may be addressed from at least two different perspectives. The first, deals with how a model would perform given external disturbances, such as errors in measurements of the variables as well as faults in the bio-sensors. The second, addressed in this section, explores how effective a model is when implemented on a different subject from the one whose data were used to develop the model.

The results shown in the paper so far were obtained by using the subject's own data to develop an initial model, which is subsequently used to predict the subject's future temperature. In practice, however, it may be necessary to provide predictions about a subject's future core temperature values without any past data from that specific subject. For this reason, we are interested in determining the effectiveness of using a model developed using data from one subject, in providing predictions of core temperature for another subject. This would provide us with a measure of robustness of the models developed in this investigation.

As previously described, an ARX model for each subject i, ($i = 1, 2, \ldots, 9$), was developed by using subject i's first half of the data from each of the two environmental conditions and subsequently tested using the second half of the data. Here, however, eight ARX models are developed for each subject i by using the first half of the data from each of the other eight subjects ($j : j \neq i, j = 1, 2, \ldots, 9$). Each of the eight models is subsequently tested using the second half of the data for each subject i. The average measure of fit over the eight hybrid models is shown in the fourth column of Tables 1 and 2. It is worth mentioning that while the ARX-part of the hybrid model was developed using data from subject j and tested on subject i, the SCENARIO-part of the hybrid model always employed input data from subject i.

5. CONCLUSIONS AND FUTURE WORK

The work presented in this paper shows that a hybrid approach (incorporating both first-principle and black-box models), coupled with closed-loop feedback of prediction errors, improves the prediction accuracy of core temperature for individual subjects. These improvements may improve the average prediction accuracy by up to 30% for prediction horizons of 20 minutes. It is also pointed out that these improvements could be more significant with the availability of more training data. In addition, the utility of a hybrid model derived from one subject's data, in predicting the core temperature of a different subject is investigated. The results show no significant improvement over using a first-principles approach by itself.

While the results shown are promising, there still remain many issues to be addressed. Chiefly, the serial hybrid approach to modeling remains to be explored.[18] In this case, a judicious application of the black-box model in improving the estimates of the parameters used in the SCENARIO model would be implemented. Special care must be taken, however, not to mask the insight provided by the SCENARIO model alone.

The issue of robustness is critical in any modeling application. Therefore, in this context, while we have investigated this question from the point of view of reusability of models, it is equally important to have some guarantees of performance when, for example, there is a slight disruption that prevents the measurement of one or more of the variables used in the models.

As is evident from the results presented in this paper, the application of the hybrid approach deteriorates as the prediction horizon increases. In order for decisions to be made on the basis of these predictions, it is critical that these predictions be reliable. Hence, it would be in order to quantitatively assess the reliability of the model predictions through estimation of statistical error bounds. One approach that will be pursued in the near future is to apply the statistical bootstrap method.[19]

While the measurements of body temperature used in the study employed in this paper involved rectal temperature measurements, future research will also investigate the effectiveness of using variables, such as skin temperature, as alternatives to core temperature, which can be measured in a non-invasive fashion to monitor and predict heat strain.

The results presented in this paper have promising implications for many other applications in the arena of physiological modeling, where there is a need to provide predictions that are subject-specific.[18]

ACKNOWLEDGMENTS

The authors would like to acknowledge the support of the following people: Miyo Yokota for highly appreciated help with the SCENARIO simulation; William Santee for the laboratory study data; Larry Berglund and Thomas McKenna for many insightful discussions; and Peg Kolka for useful comments and suggestions on the manuscript.

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. The protocol for the laboratory study[16] was funded, in part, by the Natick Soldier System Center (NSC).

DISCLAIMER

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.

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.

"This paper has been approved for public release; distribution is unlimited."

REFERENCES

1. "Heat-related injuries,us army, 2002," *MSMR Medical Surveillance Monthly Report* **9**, May/June 2003.
2. R. Hoyt and K. Friedl, "Current status of field applications of physiological monitoring for the dismounted soldier," in *Metabolic Monitoring Technologies for Military Field Applications*, M. Poos, ed., pp. 247–257, National Academy of Sciences, National Academy Press, 2004.
3. R. Hoyt, J. Reifman, T. Coster, and M. Buller, "Combat medical informatics: present and future," *AMIA 2002 Annual Symposium, San Antonio, Texas*, pp. 335–339, 2002.
4. L. Berglund and J. Stolwijk, "The use of simulation models of human thermoregulation in assessing acceptability of complex dynamic thermal environments," in *Energy conservation strategies in buildings*, J. Stolwijk, ed., pp. 157–191, Yale Press, New Haven, CT, 1978.
5. X. Xu and J. Werner, "A dynamic model of the human/clothing/environment -system," *Appl. Human Sci.* **16**, pp. 61–75, 1997.
6. M. Sawka, C. Wenger, and K. Pandolf, "Thermoregulatory responses to acute exercise-heat stress acclimation," in *Handbook of Physiology,Section 4: Environmental Physiology*, C. Blatteis and M. Fregley, eds., pp. Chapter 9,157–185, Oxford University Press for the American Physiological Society, 1996.
7. J. Stoljwijk and J. Hardy, "Control of body temperature," in *Handbook of physiology. Reactions to environmental agents*, L. D.H.K., ed., pp. 45–68, American Physiological Society, 1977.
8. K. Kraning and R. Gonzalez, "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.* **22**, pp. 331–42, 1997.
9. L. Ljung, *System Identification - Theory for the User*, Prentice Hall, New Jersey., 1999.
10. G. Goodwin and K. Sin, *Adaptive Filtering, Prediction and Control*, Prentice Hall, New Jersey, 1984.
11. R. Gonzalez, T. McLellan, W. Withey, S. Chang, and K. Pandolf, "Heat strain models applicable for protective clothing systems: Comparison of core temperature response," *J. Appl. Physiol.* **83**, pp. 1017–32, 1997.
12. R. Vilim, H. Garcia, and F. Chen, "An identification scheme combining first principle knowledge, neural networks and the likelihood function," *IEEE Transactions on Control Systems Technology* **9**, pp. 186–99, 2001.
13. J. Wilson and F. Zorzetto, "A generalized approach to process state estimation using hybrid artificial neural networks/mechanistic models," in *Computers in Chemical Engineering*, **21**, pp. 951–63, 1997.
14. S. Haykin, *Neural Networks: A Comprehensive Foundation*, Prentice Hall, New Jersey, 1998.
15. H. Demuth and M. Beale, *Neural Network Toolbox: For Use with Matlab, User's Guide, Version 4*, The MathWorks, Inc, Natick, MA, 1992.
16. W. Santee, L. Berglund, A. Cardello, C. Winterhalter, and T. L. Endrusick, "Physiological and psychological assessment of volunteers wearing air permeable battle-dress (bdu) uniforms during intermittent exercise," tech. rep., U.S. Army Research Institute of Environmental Medicine. Technical Report (in preparation).
17. N. Oleng', J. Reifman, L. Berglund, and R. Hoyt, "Alternative approaches to improve physiologic predictions," *24th Army Science Conference, Orlando, FL*, **KO-05**, 2004.
18. J. Reifman, "Alternative methods for modeling fatigue and performance," *Aviation, Space and Environmental Medicine* **75**, pp. 173–180, 2004.
19. B. Efrom and R. Tibshirani, *An Introduction to the Bootstrap*, Chapman and Hall, New York, NY, 1993.