

Estimation of Human Internal Temperature from Wearable Physiological Sensors

Mark. J. Buller^{1,2}

William J. Tharion²

Reed W. Hoyt²

Odest Chadwicke Jenkins¹

¹Dept. of Computer Science
Brown University
Providence, RI 02912
mbuller@cs.brown.edu
mark.j.buller@us.army.mil

²U.S. Army Research Inst. of
Environmental Medicine
Natick, MA 01760
william.tharion@us.army.mil

U.S. Army Research Inst. of
Environmental Medicine
Natick, MA 01760
reed.hoyt@us.army.mil

Dept. of Computer Science
Brown University
Providence, RI 02912
cjenkins@cs.brown.edu

Abstract

We evaluated a Kalman filter (KF) approach to modeling the physiology of internal temperature viewed through “noisy” non-invasive observations of heart rate. Human core body temperature (T_{core}) is an important measure of thermal state, e.g., hypo- or hyperthermia, but is difficult to measure using non-invasive wearable sensors. We estimated parameters for a discrete KF model from data collected during several Military training events and from distance runners ($n=38$). Model performance was evaluated in 25 physically-active subjects who participated in various laboratory and field studies involving exercise of 2-to-8 h duration at ambient temperatures of 20 to 40°C. Overall, the KF model’s estimate of T_{core} had a root mean square error of 0.30 ± 0.13 °C from the observed T_{core} , and was within ± 0.5 °C over 85% of the time. The benefit of the KF approach is that it requires only one input while current state of the art models typically require multiple inputs including individual anthropometrics, metabolic rate, clothing characteristics, and environmental conditions. This state estimation problem in *computational physiology* illustrates the potential for collaboration between the artificial intelligence and ambulatory physiological monitoring communities.

Introduction

In this paper, we address the problem of estimating human core temperature from non-invasive wearable sensors. This effort illustrates how an open problem in *computational physiology* can be addressed through synergistic interactions of the artificial intelligence and ambulatory physiological monitoring communities. The primary goal of human health state monitoring is to identify when an individual moves from a healthy state to a compromised state. For the elderly or individuals with certain diseases, everyday living can pose situations that can quite quickly compromise their health. For example, with diabetes, certain changes in physical activity or dietary patterns can lead to life-threatening bouts of hypo or hyperglycemia. Similarly, for healthy individuals, heavy exertion in extremes of climate or when bundled in protective clothing (e.g. see Figure 1) may quickly lead to heat illness. Predicting

health state by tracking physiological metrics such as blood glucose level or internal body temperature can help avoid dangerous and potentially life-threatening situations.



Figure 1: U.S. National Guard Civil Support Team (CST) member engaged in a chemical biological training event. The picture was taken during physiological data collection (Buller et al. 2007). The real time physiology was used by the team’s medical officer to help discern which team members were exceeding safe thermal limits. Ensuring encapsulated personnel do not overheat is difficult without measuring or estimating internal body temperature.

Today, wearable physiological sensing devices, as surveyed by (Pantlopoulos and Bourbakis 2008), have been demonstrated in a wide variety of applications including home health care, elderly monitoring, physical training, industrial hygiene, and military uses. These systems have been able to provide non-invasive physiological measurements such as heart rate, respiration rate, movement from accelerometry, galvanic skin response, and skin surface temperature to name a few. While these measures can provide a wealth of physiological information, they are quite different from what a physician might measure in a clinical setting where more controlled or invasive techniques are

available. Often the measures from the non-invasive sensors are not specific to most health states of interest and can reflect the output of multiple latent variables. Consequently, the relationship of typical ambulatory monitoring data to standard health state metrics of interest is often tenuous. In our case, for example, internal body temperature is the key to understanding hyperthermia, but non-invasive measures such as heart rate and skin temperature are also driven by the environmental conditions, clothing characteristics, individual characteristics (e.g. % body fat), work rate, hydration state, and thermoregulatory factors, and have a complex relationship to internal temperature. It is only when thermoregulation begins to fail that heart rate and skin temperature correlate more directly with internal temperature.

From the perspective of artificial intelligence, many of the ambulatory monitoring health problems can be cast as state estimation problems, where some latent internal variable is sensed indirectly from a set of “noisy” observations. In this paper we evaluate a Kalman filter (Kalman 1960) approach to model the physiology of internal temperature viewed through “noisy” observations of non-invasive variables. We demonstrate that the Kalman filter technique combined with an understanding of physiology can produce an estimator that performs at least as well as current but significantly more complex models.

We hope that the results from our experiment can serve as a baseline for physiological health state estimation research in the artificial intelligence community. Past events such as the AAAI Human Behavior Modeling Symposium (http://www.cs.dartmouth.edu/~tanzeem/hbm_ss09/program.html) and the 2004 ICML Physiological Data Modeling Contest (PDMC) (<http://www.cs.utexas.edu/users/sherstov/pdmc/>) have highlighted the challenge of working with data from ambulatory physiological monitoring systems.

Physiological data offers many challenges to the machine learning community including dealing with large amounts of data, sequential data, issues of sensor fusion, and a rich domain complete with noise, hidden variables, and significant effects of context. (PDMC Organizers)

We have collected several rich physiological field data sets that exhibit similar challenges. To foster interdisciplinary *computational physiology* research we need to overcome the institutional and practical hurdles of sharing data with the artificial intelligence community, with the goal of solving several long standing health state estimation problems.

Motivating Example

Hot environments pose a risk of heat illness for occupations where heavy workloads and/or protective clothing ensembles are necessary. Excessive heat strain can lead to collapse or even death from heat stroke (Bouchama and Knochel 2002). In a team setting these risks can be compounded by psychological pressure to continue working even if the individual is aware of feeling ill (Porter 2000). The ability to monitor core body temperature (Tcore) would provide an effective safeguard against potential heat illness for these workers.

However, classic methods for monitoring Tcore, such as rectal thermistors and esophageal probes, are not appropriate for an ambulatory setting (Gunga et al. 2008). Furthermore, Tcore estimation by heat transfer models based upon metabolic heat production is complex; an example is the initial capability decision aid (ICDA) (Yokota et al. 2008). These types of models, in addition to metabolic rate, require multiple input parameters relating to the environment, individual characteristics, and clothing.

By framing the problem in terms of a simple Bayes filter (Figure 2) we were able to identify the likely physiological parameters needed to estimate internal temperature.

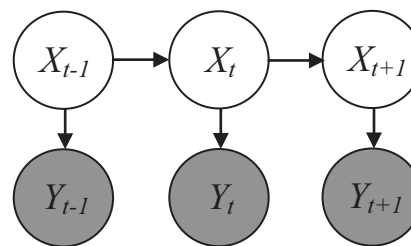


Figure 2: Graphical model showing how the typical health state estimation problem can be represented by a Bayes filter, where unobserved health states are represented by (X) and observed physiological measurements are represented by the vector (Y).

From physiology we know that as we begin to perform more demanding physical work that the body’s muscles require more oxygen. The heart responds to this increased demand with a proportional increase in heart rate (HR). The increase in work also leads to an increase in metabolic heat production. As heat is produced, it is either stored internally, causing a rise in Tcore, or is lost to the surrounding environment. Tcore is tightly regulated by the human body, usually to within 3 to 4 degrees Celsius. As work continues, and heat is stored, the body will attempt to lose heat to the environment by shunting warm blood to the skin, where it is cooled by convection or evaporative mechanisms. As the need for cooling increases, blood flow to the skin increases. Increased blood flow to the skin leads to an increase in HR beyond that needed to supply tissues with oxygen for work (Sawka and Young 2006). Our hypothesis assumes that

HR is related to both the amount of heat generated from physical work and the amount of heat being dissipated by the thermoregulatory system, suggesting that change in Tcore can be estimated by a function of HR thus:

$$\Delta T_{core} \approx \text{Heat_Gain} - \text{Heat_Loss} \approx f(\text{HR}) \quad (1)$$

In this scheme, HR is an observation that is a direct consequence of the unobserved change in Tcore, which matches the graphical model presented in Figure 2. To evaluate our hypothesis, we used data collected from field studies of soldiers and athletes to estimate the parameters of a Kalman filter.

Kalman Filter Model Development

A Kalman filter (KF) is an iterative state estimation algorithm used for linear dynamical systems with discrete time steps. To estimate a current state, the KF requires some knowledge of the previous state, the dynamics of the system, and a “noisy” measurement of the current state – in our model HR.

The KF parameters were derived from data taken from three field studies (number of test volunteers (n = 38) with over 20,000 HR-Tcore pairs (Ely et al. 2009, Hoyt et al. 2001, Tharion et al. In Preparation). We implemented a discrete KF algorithm as detailed by Welch and Bishop (1995) and is outlined in equations (2) through (7) below. At each time step, the KF uses a time update and measurement update to provide a current estimate of Tcore and the error variance. The time update uses the last estimate of Tcore and its known dynamics to predict a current value thus:

$$\hat{x}'_t = G\hat{x}_{t-1} + u \quad (2)$$

where: \hat{x}'_t = Tcore prediction for the current time t (1 min steps), the matrix G predicts a current Tcore from the previous \hat{x}_{t-1} , and u represents zero mean Gaussian noise of the prediction.

The measurement update phase uses a “noisy” measurement of the true Tcore to refine the prediction value. Initially, a measurement residual is calculated between the observation and the prediction from equation (2) mapped into the measurement space, as shown in equation (3):

$$\tilde{y}_t = H\hat{x}'_t - z_t + v \quad (3)$$

where \tilde{y} is the measurement residual, the matrix H maps the current Tcore prediction \hat{x}'_t into the HR measurement space, z_t is the current measurement of HR, and v represents zero mean Gaussian noise of the measurement. The final estimated Tcore for this time step is then computed as the predicted Tcore value from equation (2)

modified by a weighted measurement residual from equation (3), thus:

$$\hat{x}_t = \hat{x}'_t + K\tilde{y}_t \quad (4)$$

where \hat{x}_t is the current estimate of Tcore, and K is the weighting factor or Kalman gain that modifies the measurement residual to provide the current Tcore estimate. The optimal Kalman gain for the discrete KF is calculated based upon a ratio of the error variances of u and v and the error variance from the previous time step. The error variance for the time update is estimated as follows:

$$P'_t = GP_{t-1}G^T + Q \quad (5)$$

where P'_t is the time update error variance, P_{t-1} is the error variance for the previous Tcore and Q is the time update error variance based upon u .

The Kalman gain K is then calculated as follows:

$$K_t = P'_t H^T (H P'_t H^T + R)^{-1} \quad (6)$$

where R is the measurement update error variance based upon v . Finally the current Tcore estimate error variance is computed:

$$P_t = (I - K_t H) P'_t \quad (7)$$

To define our KF algorithm, the mapping matrices G and H and the noise components u and v were estimated from our training data. Because of thermal inertia (Gribok et al. 2008), the matrix G was defined as an identity matrix that predicts the same Tcore as estimated at the previous time step.

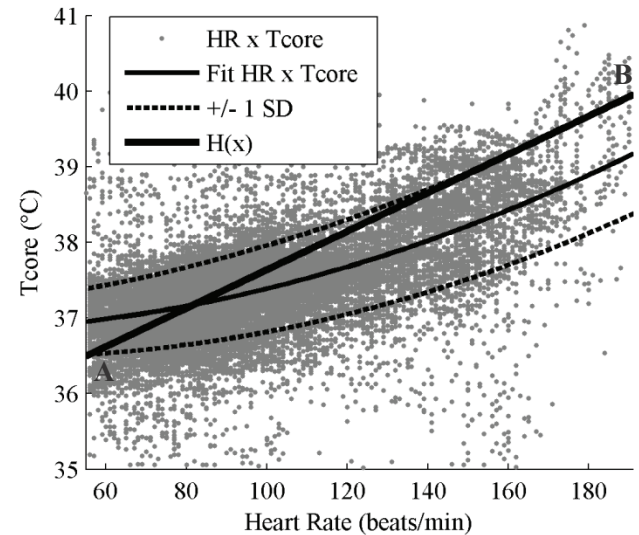


Figure 3: Training Data for the Kalman Filter mapping matrix H . Scatter plot of HR (binned by 1 beat/min) by Tcore with quadratic least squares fit of mean Tcore \pm 1 SD, and the KF mapping matrix function $H(x)$ drawn between points A and B.

A linear relationship between Tcore to HR was derived from the data (see Figure 3). We assumed that using a regression function on our HR and Tcore training data would result in a mapping function where the gradient term was too small, since for any given HR, Tcore should, in time, approach a preferred value. Since our training data is dynamic in nature (i.e., HR and Tcore were variable and did not often reach a steady state), we found that there were many Tcore values for any given HR bin. Therefore, we assumed that for low values of HR, the preferred value of Tcore would be among the lower Tcore values, and for high HRs, the preferred value would be among the higher Tcore values. Using this assumption, we defined H as a line drawn between a low HR value and the mean Tcore $- 1$ SD (A), and a high HR and the mean Tcore $+ 1$ SD (B). Where A = (56 beats/min, 36.5 °C) and B = (190 beats/min, 39.9 °C). H is defined as the equation of the line drawn between the points A and B, thus:

$$H(x) = 39.3701x - 1381.6890 \quad (8)$$

The discrete probability distribution of the change in Tcore from minute to minute was learned directly from the data and is presented in Figure 4. By correlation we determined that change in core temperature varies independently of the actual core temperature value ($r = 0.016$). The time update error variance for Q was computed from the change in Tcore probability distribution: $Q = 0.000576$.

Over all discrete core temperatures the median standard deviation was 18 beats/min., thus the measurement update error variance was defined as: $R = 324$.

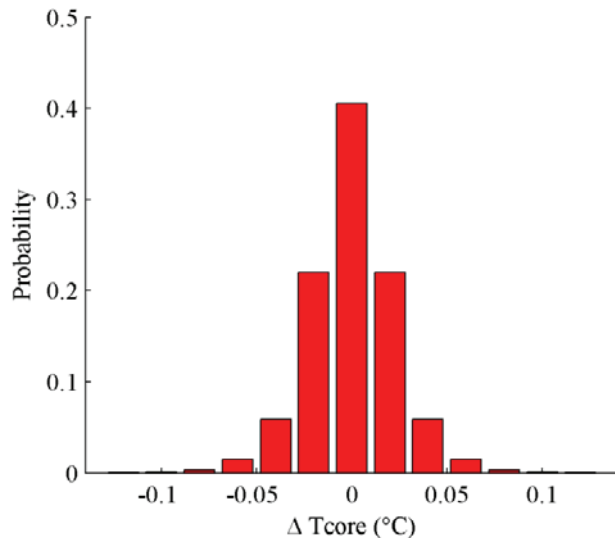


Figure 4: Discrete probability distribution for change in core body temperature.

Kalman Filter Model Evaluation

The performance of our KF model was examined using a separate set of data from three studies ($N = 25$), with $>25,000$ HR-Tcore pairs (Cheuvront et al. 2007, Hoyt et al. 2002, Hoyt and Friedl 2004). The studies were in both laboratory and field settings and provided varied environmental conditions (20°C to 40°C), low to moderate work rates, and moderate to long durations (2 to 8 hours). The studies included a series of six eight-hour work-rest regimes conducted within an environmental chamber, a two hour military field training, and data from a military cadet suffering from heat exhaustion.

Comparisons between the KF model estimate and observed Tcore were examined overall and by each experimental condition. Summary statistics of root mean square deviation (RMSD), bias, and non-parametric Bland-Altman percentage (BA) (Bland and Altman 1999) were calculated. The BA calculates the percentage of estimated Tcore points falling within an a priori zone of the actual Tcore ($\pm 0.5^\circ\text{C}$ was used in our analysis). Additionally the performance of our KF model was compared to the more complex ICDA model.

Results

Overall, Tcore estimated by the KF model had a RMSD of 0.30 ± 0.13 °C from the observed Tcore. The KF had a bias of 0.02 ± 0.18 °C, and over 85% of all estimated Tcore data points were within $\pm 0.5^\circ\text{C}$ of the observed value.

Table 1 shows the performance of the KF model and the ICDA model under eight different experimental conditions. In 7 of 8 conditions the KF model provided a closer estimation of Tcore with lower variance.

Condition	N	KF Model		ICDA
		RMSD (°C)	BA (%)	RMSD (°C)
L: 20°C/wr=M	9	0.32±0.14	85.7	0.35±0.34 [†]
L: 27°C/wr=L	11	0.27±0.10	92.5	0.31±0.27 [†]
L: 27°C/wr=M	10	0.38±0.14	74.7	0.34±0.35 [†]
L: 35°C/wr=L	12	0.28±0.14	87.5	0.29±0.31 [†]
L: 35°C/wr=M	7	0.32±0.14	87.6	0.37±0.41 [†]
L: 40°C/wr=L	7	0.31±0.14	82.9	3.31±0.62 ^{†‡}
F: 25°C/wr=I	6	0.26±0.13	89.7	0.42±0.24
F: 32°C/wr=I	1	0.17	94.5	0.69
Mean ± SD		0.30±0.13	85.9	0.34±0.31

Table 1: KF and ICDA model performance for laboratory (L:) and field (F:) settings. Work Rate (wr) is = light (L, $\sim 175\text{W/m}^2$), moderate (M, $\sim 230\text{W/m}^2$), or intermittent (I). [†] Degroot et al. 2008. The highlighted ([‡]) value exceeds ICDA assumptions, and is not included in the overall weighted mean.

Figure 5 shows the KF estimated Tcore and observed Tcore for an individual military cadet who became a heat exhaustion casualty during a military training event. For this individual the model has an average RMSD of 0.17°C with over 94% of the estimates within $\pm 0.5^\circ\text{C}$.

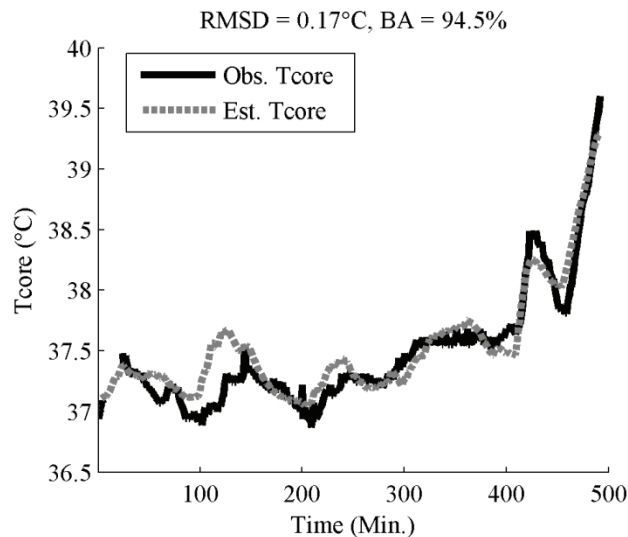


Figure 5: Observed (obs.) and estimated (est.) Tcore responses for an individual military cadet who became a heat exhaustion casualty during training. RMSD = root mean square deviation, and BA = non-parametric Bland Altman percentage.

Discussion and Conclusions

A rich set of ambulatory physiological data allowed us to generate the parameters for a Kalman filter (KF) to estimate core body temperature (Tcore). The KF estimator worked at least as well as a current physiological model (Yokota et al., 2008). The benefit of the KF approach is that it requires only one input – heart rate (HR) – while the current state-of-the-art models require a multitude of inputs including metabolic rate, environment conditions, clothing characteristics, and individual anthropometrics.

In this paper, the KF model was constructed in a simple manner. First, the prediction function $G(X)$ was an identity operator predicting that the new Tcore was the same as the old Tcore. Second, the mapping function $H(X)$ was assumed to be linear, which at the extremes (approximately greater than 39.5°C and less than 36.5°C) will not be the case. For example, humans have a maximum heart rate, and as HR approaches this, Tcore can continue to rise while HR remains constant. Finally, we assumed that variances for the prediction and mapping functions are constant, which at least for the mapping function is unknown.

While the model was simple, physiological knowledge was used to formulate the problem in terms of a latent variable that could be inferred from a “noisy” observation. The combination of physiological insight to

formulate the problem and the application of a simple state estimation technique has achieved a model that performs as well as its peers in the physiology community.

Although the current construction of the KF will not correctly model extreme core temperatures above 39.5°C we suggest that with a few refinements this model could offer a promising new means of real-time thermal injury risk assessment. We suggest that the other parameters of human health state monitoring are ripe for the intelligent application of machine learning techniques. For the military, performance athletes, first responders, etc., health state estimators hydration state, metabolic state, fatigue, and cognitive state (see Friedl 2004) remain elusive using non-invasive sensors. From a military combat casualty care perspective, lives can be saved if remote monitoring can detect loss of blood (Rickards et al. 2008). Similarly, from a diabetes management point of view, the ability to estimate and guard against hypoglycemic shock is a critical step on the path to an artificial pancreas (Hoshino et al. 2009).

The need to share non-invasive ambulatory physiological and contextual data collected during complex, real-world military training exercises with the artificial intelligence community is apparent. Analysis using innovative methods developed by the artificial intelligence community will likely lead to new metrics and markers of physiological status. Yet, the sharing of data requires the proper safeguards regarding data integrity, confidentiality of the data, and proper recognition in terms of authorship of the investigators that generated the data. These hurdles can be addressed through proper data sharing agreements to advance this line of research.

We conclude that *computational physiology* has the potential to solve many difficult ambulatory monitoring problems that can be used in real world applications to prevent the onset of illness or prevent death.

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 Army or the Department of Defense. The investigators have 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 32 CFR Part 219. Citations of commercial organizations and trade names in this report do not constitute an official Department of the Army endorsement or approval of the products or services of these organizations.

References

- Bland, J. M., and Altman, D. G. 1999. Measuring agreement in method comparison studies. *Statistics Methods in Medical Research* 8: 135-160.
- Bouchama, A., and Knochel, J. P. 2002. Heat Stroke. *N. Engl. J. Med.* 346(25): 1978-1988.
- Buller, M. J.; Tharion, W. J.; Karis, A. J.; Santee, W. R.; Mullen, S. P.; Blanchard, L. A.; and Hoyt, R. W. 2007. Demonstration of Real-Time Physiological Status Monitoring of Encapsulated 1st Civil Support Team-Weapons of Mass Destruction (CST-WMD) Personnel, Technical Report. T08-01/ADA 473 188, USARIEM, Natick, MA.
- Cheuvront, S. N.; Montain, S. J.; Goodman, D. A.; Blanchard, L.; Sawka, M. N. 2007. Evaluation of the limits to accurate sweat loss prediction during prolonged exercise. *Eur. J. Appl. Physiol.* 101(2): 215-224.
- DeGroot, D. W.; Goodman, D. A.; Montain, S. J.; and Cheuvront, S. N. 2008. Validation of the ICDA Model for Predicting Body Core Temperature. *Medicine & Science in Sports & Exercise.* 40(5): S367 (Abstract).
- Ely, B. R.; Ely, M. R.; Cheuvront, S. N.; Kenefick, R. W.; DeGroot, D. W.; and Montain, S. J. 2009. Evidence against a 40°C core temperature threshold for fatigue in humans. *J. Appl. Physiol.* 107:1519-1525.
- Friedl, K. E. 2004. Predicting and protecting performance using metabolic monitoring strategies: it's all in the wet stuff anyway, isn't it? In: *Monitoring Metabolic Status – Predicting Decrements in Physiology and Cognitive Performance.* Washington, DC: The National Academies Press.
- Gribok, A. V.; Buller, M. J.; and Reifman, J. 2008. Individualized short-term core temperature prediction in humans using biomathematical models. *IEEE Transactions on Biomedical Engineering.* 55(5), 1477-1487.
- Gunga, H. C.; Sandsund, M.; Reinertsen, R. E.; Sattler, F.; and Koch, J. 2008. A non-invasive device to continuously determine heat strain in humans. *J. Thermal Biology.* 33: 297-307.
- Hoshino, M.; Haraguchi, Y.; Mizushima, I.; and Sakai M. 2009. Recent progress in mechanical artificial pancreas. *Journal of Artificial Organs.* 12(3): 141-149.
- Hoyt, R. W.; Buller, M. J.; DeLany, J. P.; Stultz, D.; Warren, K.; Hamlet, M. P.; Shantz, D.; Matthew, W. T.; Tharion, W. J.; Smith, P.; and Smith, B. 2001. Warfighter physiological status monitoring (WPSM): energy balance and thermal status during a 10-day cold weather U.S. Marine Corps Infantry Officer Course field exercise, Technical Report, T02-02/ADA396133, USARIEM, Natick, MA.
- Hoyt, R. W., and Friedl, K. E. 2004. Current status of field applications of physiological monitoring for the dismounted soldier. In: *Monitoring Metabolic Status – Predicting Decrements in Physiology and Cognitive Performance.* Washington, DC: The National Academies Press.
- Hoyt, R. W.; Reifman, J.; Coster, T. S.; and Buller, M. J. 2002. Combat medical informatics: present and future. In: *Biomedical Informatics: One Discipline*, edited by: Kohane, I. S. Proceedings of the AMIA Annual Symposium ISBN: 1-5603-600-4.
- Kalman, R. E. 1960. A new approach to linear filtering and prediction problems. *Journal of Basic Engineering* 82(1): 35–45.
- Pantlopoulos, A., and Bourbakis, N. 2008. A Survey on Wearable Biosensor Systems for Health Monitoring. In *30th Annual International IEEE Engineering in Medicine and Biology Society (EMBS) Conference.* 4887-4890.
- Porter, A. M. W. 2000. The death of a British officer-cadet from heat illness. *Lancet.* 355: 569-571.
- Rickards, C. A.; Ryan, K. L.; Cooke, W. H.; Romero, S. A.; and Convertino, V. A. 2008. Combat stress or hemorrhage? Evidence for a decision-assist algorithm for remote triage. *Aviat. Space Environ. Med.* 79:670-676.
- Sawka, M. N. and Young, A. J. 2006. Physiological systems and their responses to conditions of heat and cold. In: *ACSM's Advanced Exercise Physiology*, edited by Tipton, C. M.; Sawka, M. N.; Tate, C. A.; Terjung, R. L. American College of Sports Medicine, Lippincott, Williams & Wilkins. New York, NY.
- Tharion, W. J.; Karis, A. J.; Buller, M. J.; Patel, T.; Yokota, M.; Blanchard, L.; Santee, W. R.; Mullen, S. P.; Hebert, N. J.; and Hoyt, R. W. (In Preparation). Accuracy of the Heat Strain Decision Aid (HSDA) during Ranger Training Brigade's road march training exercise, Technical Report, T10-XX, Natick, MA.
- Welch, G., and Bishop G., 1995. An Introduction to the Kalman Filter, Technical Report, TR 95-041, Department of Computer Science, University of North Carolina at Chapel Hill.
- Yokota, M.; Berglund, L.; Cheuvront, S.; Santee, W.; Latzka, W.; Montain, S.; Kolka, M.; and Moran, D. 2008. Thermoregulatory model to predict physiological status from ambient environment and heart rate. *Computers in Biology and Medicine.* 38: 1187-1193.