Estimation of human core temperature from sequential heart rate observations

11 authors, including:

Mark Buller
US Army
101 PUBLICATIONS   **1,113** CITATIONS

SEE PROFILE

William J. Tharion
U.S. Army Research Institute of Environmental Medicine
122 PUBLICATIONS   **2,272** CITATIONS

SEE PROFILE

Scott Montain
U.S. Army Research Institute of Environmental Medicine
249 PUBLICATIONS   **12,177** CITATIONS

SEE PROFILE

Robert W Kenefick
United States Army
162 PUBLICATIONS   **5,176** CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Cardiovascular Health View project

Can normobaric hypoxia maintain hypobaric hypoxia acclimatization? View project

Estimation of human core temperature from sequential heart rate observations

Mark J Buller[1,4]**, William J Tharion**[1]**, Samuel N Cheuvront**[1]**,
Scott J Montain**[1]**, Robert W Kenefick**[1]**, John Castellani**[1]**,
William A Latzka**[1]**, Warren S Roberts**[2]**, Mark Richter**[3]**,
Odest Chadwicke Jenkins**[4] **and Reed W Hoyt**[1]

[1] US Army Research Institute of Environmental Medicine, Natick, MA 01760, USA
[2] Defence Science and Technology Organization, Melbourne, Australia
[3] Marine Corps System Command, Quantico, VA 22135, USA
[4] Department of Computer Science, Brown University, Providence, RI 02912, USA

E-mail: mark.j.buller.civ@mail.mil

Abstract

Core temperature (CT) in combination with heart rate (HR) can be a good
indicator of impending heat exhaustion for occupations involving exposure to
heat, heavy workloads, and wearing protective clothing. However, continuously
measuring CT in an ambulatory environment is difficult. To address this
problem we developed a model to estimate the time course of CT using a series
of HR measurements as a leading indicator using a Kalman filter. The model was
trained using data from 17 volunteers engaged in a 24 h military field exercise
(air temperatures 24–36 °C, and 42%–97% relative humidity and CTs ranging
from 36.0–40.0 °C). Validation data from laboratory and field studies ($N = 83$)
encompassing various combinations of temperature, hydration, clothing, and
acclimation state were examined using the Bland–Altman limits of agreement
(LoA) method. We found our model had an overall bias of -0.03 ± 0.32 °C and
that 95% of all CT estimates fall within ± 0.63 °C ($>52\,000$ total observations).
While the model for estimating CT is not a replacement for direct measurement
of CT (literature comparisons of esophageal and rectal methods average LoAs
of ± 0.58 °C) our results suggest it is accurate enough to provide practical
indication of thermal work strain for use in the work place.

Keywords: core body temperature, Kalman filter model, latent variable, core
temperature surrogate, non-invasive sensors

1. Introduction

Continuous ambulatory measurement of core body temperature (CT) can be a critical component of human heat strain assessment during strenuous work (Moran *et al* 1998; Frank *et al* 2001). However, while personal physiological monitoring technology has developed to the point where multi-parameter sensor systems can be used to collect data in a variety of settings over extended periods of time, the requisite measurement of CT still remains a challenge.

Medical grade CT measurement using pulmonary arterial blood temperature is only appropriate in a clinical setting. The traditionally accepted laboratory rectal and esophageal probe methods are impractical for ambulatory settings. Other non-invasive methods of estimating CT using external measurements such as axillary or tympanic temperatures have proven unreliable (Lim *et al* 2008). Ingestible thermometer pills (e.g., Jonah Pill thermometer, Respironics, Bend, OR) have been used successfully in field settings (e.g. Lee *et al* 2010), and have been within acceptable limits of agreement (LoA) (± 0.4 °C) and bias (<0.1 °C) when compared to esophageal temperatures (Byrne and Lim 2007). However, these thermometer pills have drawbacks: (1) they cannot be used by all people due to medical contraindications, and (2) can suffer from inaccuracy when hot or cold fluids are consumed (Wilkinson *et al* 2008). The difficulty in directly measuring CT in ambulatory settings has led to the search for a practical alternative technique.

One non-invasive approach that has received attention is the zero heat flux (ZHF) method (Fox *et al* 1973) where an insulated area of the skin is heated until there is no heat flow. The temperature of the skin is then assumed to be equivalent to deep body temperature. Most of the work on this approach has been in laboratory and clinical settings (Yamakage *et al* 2002) with recent work focusing on improving measurement of dynamic temperature changes (Steck *et al* 2011); decreasing the technique's response time (Teunissen *et al* 2011); and adapting the ZHF method for use in ambulatory environments (Gunga *et al* 2008 2009). In clinical settings these devices have demonstrated good agreement with esophageal measures, while custom sensors developed for ambulatory environments have had varying degrees of success depending on environmental conditions.

Other researchers have used thermoregulatory heat transfer models to estimate CT (Kraning and Gonzalez 1997, Fiala *et al* 2001, Havenith 2001). These models use an array of input variables that include metabolic rate, environmental parameters, individual characteristics, and clothing parameters (insulation and vapor permeability). In an ambulatory field setting these models suffer from the fact that not all inputs are available all of the time, and measuring or estimating metabolic rate is difficult. Recent work has focused on combining thermoregulatory heat transfer models with metabolic rate estimators that use heart rate (HR) with ambient temperature modifiers to account for skin blood flow (Yokota *et al* 2008). This real-time model provided accurate group-mean CT estimates in a number of different environmental and clothing conditions (Degroot *et al* 2008). While this method shows promise it still requires many input parameters that must be measured independently from an individual such as environmental conditions and clothing characteristics.

Concentrating on estimating CT in warm to hot conditions during exercise we present a method to use time series observations of HR to track CT over time. Our method relies on a Kalman filter (KF) (Kalman 1960) which has been used extensively in engineering tracking problems. Here an item or variable of interest must be tracked from a series of 'noisy' observations, and knowledge of the temporal dynamics. The KF requires two models defined by linear Gaussian probability density functions. One model relates how the variable to be tracked changes over time, while the other model relates current observations to the variable

of interest. We hypothesized that HR could be used as a 'noisy' observation of CT. Thus, by understanding how CT changes over time and the most likely CT for a given HR, a KF model to estimate a series of CT values could be learned. HR is a convenient observation of the expected CT at steady state or a leading indicator of CT as it contains information about both heat production (through the Fick (1855) equation and VO_2) and heat transfer since HR is related to skin profusion. In our previous work (Buller *et al* 2010) we demonstrated the feasibility of the KF method for estimating CT, but the model artificially limited CT estimation to values below 39.5 °C and lacked systematic validation using data from a variety conditions known to impact CT. This paper extends our previous KF CT estimation model in the following ways: (1) we use an extended KF (see Welch and Bishop 1995) to allow estimation of CT up to 41 °C, (2) the model's CT time update and CT to HR mapping functions are derived from a single study with 17 volunteers where CT ranges from 36 °C to over 40 °C, and (3) the model is validated against original data sets from laboratory and field experiments where work rates and environment, hydration, clothing, and acclimation states are varied. Our goal is to provide a method of estimating CT in warm-hot environments that is simple to use, works with equipment that is readily available and provides a valid estimate of time-varying CT.

2. Methods

2.1. Test volunteers

Data from ten laboratory and field studies with a total of 100 test volunteers were used in the development ($N = 17$), and validation ($N = 83$) of the KF model. Original data from the studies were used in consultation with the principal investigators. These studies are described in detail in their original cited publications. All research was conducted under the oversight of Institutional Review Boards. In some instances, the number of volunteers used for our analyses was less than those reported in the cited studies. These instances occur where either the HR and/or CT data were not available for these participants from the original research data or where volunteers failed to complete the whole experiment. Table 1 contains a summary of study volunteers, work rates, and environmental conditions.

2.1.1. Model development data. HR (Equivital EQ02, Hidalgo Cambridge UK) and CT (ingested—Jonah Thermometer Pill, Respironics, Bend OR) data were collected from 17 male US Army volunteers (age = 23 ± 4 yrs, height = 1.79 ± 0.08 m, weight = 81.3 ± 10.8 Kg, body fat = 18 ± 3% mean ± standard deviation (SD)) on one of two days of a field training exercise during July 2011 (air temperature 24–36 °C, 42–97% relative humidity (RH), wind speeds from 0 to 4 ms^{-1} with activities during the day conducted under full sun) at Fort Bragg, North Carolina. The field exercise included periods of sleep, rest, foot movement and periods of vigorous upper body work, providing a very wide range of work rates. These data were chosen to develop the model as they included the largest range (36–40 °C) and most dynamic CT responses of our analyzed data.

2.1.2. Model validation. Data from nine studies were used to examine the performance of the model in a number of different conditions. Four laboratory studies were used for controlled comparisons of the effects of different environments, hydration states, clothing ensembles, and acclimation state; and five field physiological monitoring experiments, were used to examine the performance under different climates and different levels of protective clothing.

Table 1. Volunteer characteristics, TEE rate and environment summary by study.

Study	Time (min)	n	Age (yrs)	Height (m)	Wt (kg)	Body fat (%)	TEE Rate (W)[a]	Air temp. (°C)	RH (%)
T	~840	17	23 ± 4	1.79 ± 0.08	81 ± 11	18 ± 3	Various	24–36	42–97
A	~480 × 6	18[b]	22 ± 4	1.77 ± 0.04	81 ± 15	N/C	350/470	20–40	30–50
B	121/121	8	23 ± 3	N/C	72 ± 12	N/C	1000	33	50
C	111/28	6/8	23 ± 6	1.76 ± 0.06	76 ± 15	18 ± 6	675	35	55
D	59/100	7	24 ± 7	1.78 ± 0.08	80 ± 21	16 ± 11	550	45	20
E	140	11	27 ± 6	1.77 ± 0.05	82 ± 5	14 ± 3	675	25	85
F	1441	7	27 ± 2	1.78 ± 0.08	86 ± 6	N/C	200	9–13	83–95
G	209 + 250	8	21 ± 1	1.80 ± 0.07	85 ± 9	15 ± 3	200	39–47	9–13
H	683 + 488	8	21 ± 2	1.84 ± 0.04	86 ± 6	16 ± 3	400	20	20–26
I	297/244	8	28 ± 6	1.95 ± 0.09	86 ± 14	13 ± 4	Var./685	15–20	65–85

TEE = Total energy expenditure rates.
[a] Values reported are approximate. T = Training/development data.
[b] Includes 1 female. N/C = Not collected. Var. = various. Means ± SD.

*Laboratory study (**A**) environmental conditions* (Cheuvront *et al* 2007)*:* 18 volunteers (1 female) (22 $\pm$ 4 yrs, 1.77 $\pm$ 0.04 m, 80.9 $\pm$ 15.3 kg) participated in six 8 h bouts of intermittent treadmill exercises while wearing US Army battledress uniform. Volunteers were euhydrated and heat acclimated. CT was measured using a thermometer pill suppository. The six test conditions were: (A.1) 20 °C, 50% RH and a total energy expenditure (TEE) rate of ~460 W; (A.2) 27 °C, 40% RH and a TEE rate of ~350 W; (A.3) 27 °C, 40% RH, and a TEE rate of ~470 W; (A.4) 35 °C, 30% RH and an TEE rate of ~350 W; (A.5) 35 °C, 30% RH and a TEE rate of ~470 W; and (A.6) 40 °C, 40% RH and a TEE rate of ~360 W.

*Laboratory study (**B**) hydration state* (Montain and Coyle 1992)*:* eight heat acclimated male volunteers (23 $\pm$ 3 yrs, 71.9 $\pm$ 11.6 kg) completed 2 h of cycle ergometer exercise at a TEE rate of ~1000 W while wearing shorts and a t-shirt in environmental conditions of 33 °C, 50% RH. CT was measured with a rectal probe. Conditions were: (B.1) Hydrated with 80% fluid replacement; and (B.2) dehydrated with no fluid replacement.

*Laboratory study (**C**) clothing* (Latzka *et al* 1997, 1998)*:* eight heat acclimated euhydrated male volunteers (23 $\pm$ 6 yrs, 1.76 $\pm$ 0.06 m, 76.0 $\pm$ 15.1 kg, 18 $\pm$ 6% body fat) participated in treadmill exercise at TEE rates of ~675 W in an environment of 35 °C and 55% RH. CT was measured with a rectal probe. Conditions were: (C.1) shorts and a t-shirt (n = 6) for 111 min of exercise; and (C.2) totally encapsulating chemical protective clothing ($n = 8$) for 28 min of exercise.

*Laboratory study (**D**) acclimation state* (Kenefick *et al* 2011)*:* seven male euhydrated volunteers (24 $\pm$ 7 yrs, 1.78 $\pm$ 0.08 m, 80.2 $\pm$ 21.3 kg, 16 $\pm$ 11% body fat) participated in a treadmill exercise at a TEE rate of ~550 W while wearing shorts and a t-shirt in environmental conditions of 45 °C, 20% RH. CT was measured using a thermometer pill used as a suppository. Conditions were: (D.1) unacclimated for 59 min of exercise; and (D.2) acclimated (ten previous days of exercise in the heat) for 100 min of exercise.

*Field study (**E**) US army ranger training brigade (RTB) (unpublished):* 11 male acclimated euhydrated RTB students (27 $\pm$ 6 yrs, 1.77 $\pm$ 0.05 m, 81.7 $\pm$ 5.3 kg, 14 $\pm$ 3% body fat) participated in an 8 mile timed road march (140 min) while carrying ~ 35 kg at night. Volunteers wore the Army combat uniform, and had TEE rates of ~675 W in 25 °C, 85% RH environmental conditions with wind speeds ranging from 0 to 3 ms^{-1}. CT was measured by ingested thermometer pill.

*Field study (**F**) US special forces* (Buller *et al* 2011b)*:* seven male heat acclimated euhydrated special forces military students (27 $\pm$ 2 yrs, 1.78 $\pm$ 0.08 m, 85.7 $\pm$ 6.2 kg) who were participating in multi-day selection course were studied. Volunteers were studied over a 24 h period which included various training activities and sleep. Volunteers wore the Army combat uniform and had average TEE rates of ~200 W. Environmental conditions ranged from 9 to 13 °C and 83 to 95% RH with wind speeds of 0.4 to 3.0 ms^{-1} with some sun during outdoor activities. CT was measured by ingested thermometer pill.

*Field study (**G**) Iraq* (Buller *et al* 2008)*:* eight male heat acclimated euhydrated US Marines (21 $\pm$ 1 yrs, 1.80 $\pm$ 0.07 m, 85.1 $\pm$ 9.0 kg, 15 $\pm$ 3% body fat) who conducted one of two foot patrols (209 and 250 min) in Iraq were studied. Volunteers wore the standard Marine Corps combat shirts and body armor (~37 kg load) and had an average TEE rate of ~200W. Environmental conditions were 42 to 47 °C and 9 to 11% RH; and 39 to 44 °C, and 9 to 13% RH with wind speeds <2.0 ms^{-1}. Both patrols were conducted in full sun. CT was measured by ingested thermometer pill.

*Field study (**H**) Afghanistan* (Buller *et al* 2011a)*:* eight male heat acclimated US Marines (21 $\pm$ 2 yrs, 1.84 $\pm$ 0.04 m, 85.7 $\pm$ 6.2 kg, 16 $\pm$ 3% body fat) who conducted one of two foot patrols during a full mission day in Afghanistan were studied (683 and 488 min).

Volunteers wore the standard Marine Corps combat shirts and body armor (∼32 kg load). Patrols were conducted with average TEE rates of ∼400 W. Environmental conditions were $20 \pm 3\,°C$, and $20 \pm 11\%$ RH with wind speeds of $2.4 \pm 0.8\,\mathrm{ms}^{-1}$; and $20 \pm 5.3\,°C$, $26 \pm 13\%$ RH with wind speeds of $2.0 \pm 1.1\,\mathrm{ms}^{-1}$. Both monitoring periods were under full sun. CT was measured by ingested thermometer pill.

*Field study (**I**) Australian army soldiers (unpublished):* eight male heat acclimated euhydrated Australian army soldiers (28 ± 6 yrs, 1.95 ± 0.09 m, 85.7 ± 14.2 kg, $13 \pm 4\%$ body fat) participated in two training activities. Conditions were: (I.1) a simulated patrol and ambush (15–$20\,°C$, 65–85% RH, wind speed $<1.5\,\mathrm{ms}^{-1}$, limited sun) which included periods of strenuous activity (297 min). Volunteers wore chemical biological protective gear in an open configuration (Military Operational Protective Posture (MOPP) II). (I.2) A 5 km road march conducted in fully encapsulating chemical biological protective equipment worn in the MOPP IV configuration ($18\,°C$, 72% RH, wind speed $<1\,\mathrm{ms}^{-1}$, dusk) with an average TEE rate of ∼685 W (244 min). CT was measured by ingested thermometer pill.

2.2. KF model development

A KF model is comprised of two relationships: a *time update model,* and an *observation model.* In the estimation of CT the *time update model* relates how CT changes from time step to time step along with the uncertainty/noise of this change. The *observation model* relates an observation of HR to a CT value along with the uncertainty of this mapping. The *time update* and *observation* models are shown in equations (1) and (2) as regression models with the uncertainty/noise represented as zero mean Gaussian distributions with variances of γ^2 and σ^2. If the model parameters (a_0, a_1, γ, b_0, b_1, b_2, and σ) can be found, a standard set of KF equations can be used to iteratively compute the most likely CT given a series of HR observations (see equations (3) through (8) in the *Results* section). The equations in the results section are shown using the model parameters from equations (1) and (2), and where the learned model parameters have been substituted. Thus, given any series of one minute HR observations these equations can be used to iteratively compute a series of minute by minute CT estimates. Welch and Bishop (1995) provide an extensive tutorial on the KF and the extended KF which is used in this paper. However, at each time step the KF equations can be thought of as operating in the following way: (1) compute an estimate of the current CT using the time update model (see equation (3)). (2) Compute the uncertainty of the current CT estimate using the time update model uncertainty (see equation (4)). (3) Adjust the current CT estimate using the current observation of HR and the observation model weighted by the uncertainty of the observation versus the uncertainty of the current CT estimate (see equation (7)). (4) Adjust the CT estimate uncertainty based upon the uncertainty of the observation (see equation (8)). The KF model parameters were learned in the following way:

2.2.1. Time update model.
The time update model was defined as a linear regression equation as follows:

$$\mathrm{CT}_t = a_1\mathrm{CT}_{t-1} + a_0 + f \text{ Where} : f \sim N(0, \gamma) \tag{1}$$

Where CT = core temperature, subscript t = time point, a_1 = time update model coefficient, a_0 = time update model intercept, f = noise drawn from a Gaussian distribution (N) with mean 0 and SD γ. Parameters a_1 and a_0 were found by least squares regression of CT_t by CT_{t-1}. The parameter γ was derived from the SD of the discrete probability distribution of ΔCT points from the development data.

2.2.2. Observation model. The observation model was defined as a quadratic regression model as follows:

$$\text{HR}_t = b_2 \text{CT}_t^2 + b_1 \text{CT}_t + b_0 + g \text{ Where} : g \sim N(0, \sigma) \qquad (2)$$

Where b_2 = observation model quadratic coefficient, b_1 = observation model coefficient, b_0 = observation model intercept, g = noise drawn from a Gaussian distribution with mean 0 and SD σ. Equation (2) shows a quadratic regression model as this was found to better fit the development data necessitating the use of the extended KF. Parameters b_0, b_1, and b_2 were found by quadratic least squares regression fit to eight pairs of CT–HR points found by searching for the optimal CT estimation performance of our previous KF model (Buller *et al* 2010). The parameter σ was found by computing the mean and SD of HR values binned by CT at 0.1 °C intervals and taking the mean of the SD values for each bin.

Our original KF model (Buller *et al* 2010) was used to search for the optimal CT–HR points using our developmental data. The original KF linear observation model was split into seven line segments at eight CT values of 36.5, 37.0, 37.5, 38.0, 38.5, 39.0, 39.5, and 40.0 °C. Our original KF model was modified to be run in a piecewise fashion using these seven line segments. For each CT (listed above) starting with the lowest we systematically varied the HR value (± 50 beats/min in 1 beat intervals) to redefine the KF observation model at this point. For each HR we used the redefined KF model to provide estimates of CT given our development data. The HR that provided CT estimates with the minimum root mean square error (RMSE) compared to the observed development data was selected. The next highest CT line segment point was then selected and the process repeated. In this way the eight CT–HR pairs were modified by our developmental data from our earlier observation model to a new model that better defined the relationship between CT and HR. A quadratic least-squares regression was fit to these points to become our optimized observation model.

2.3. Model validation (statistical analysis)

The LoA method (Bland and Altman 1986) was selected as the most appropriate means for assessing agreement between the observed CT and KF model estimate. This method plots the average of observed and estimated values against the difference (estimate – observation). Bias is computed as the mean of the differences. LoA are computed as bias $\pm 1.96 \times$ SD of the differences. The LoA provide a range of error within which 95% of all estimates using the KF approach should fall assuming a normal distribution. The initial observed CT values for each study were used as starting values for the KF model and the initial variance was set to zero indicating high confidence in these values. The KF model was developed using data with one minute intervals. Where data had sampling rates more frequent than the one min intervals the mean of all values occurring in that minute was used. Where the sampling rate was greater than one min, values were linearly interpolated.

The Bland and Altman method specifies no *a priori* limits on what forms an acceptable bias or range of LoA; instead they suggest these values depend on the measure and its intended use. For this analysis, we compared our model's performance to how the accepted laboratory measures of rectal and esophageal temperatures compare. Bias limits were set to the individual biological variation of ± 0.25 °C found by Consolazio *et al* (1963). To set LoA we examined the literature for comparisons of rectal to esophageal temperatures. Table 2 shows results of bias $\pm$ SD and LoA for five studies. Taking a weighted mean of all studies suggests that 95% of comparisons of rectal versus esophageal CTs fall within $\pm$ 0.58 °C. This LoA appears to reflect the difficulty in obtaining tight agreement in different methods of CT measurement. This difficulty is highlighted when both esophageal and rectal temperature methods are compared
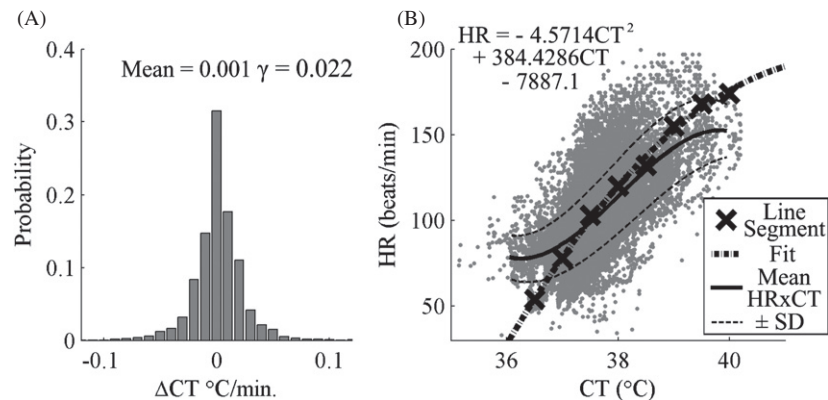
Figure 1. (A) *Time update* model represented as a discrete probability distribution found from the development data. (B) *Observation* model. Scatter plot of development data points showing mean HR by CT $\pm$ SD the optimal CT–HR line segment points (line segment) and the CT to HR mapping function (fit).

Table 2. Bias and LoA for studies comparing rectal and esophageal measure of CT.

Citation	Bias $\pm$ SD	LoA (1.96 * SD)	N
Kolka[a]	–0.21 $\pm$ 0.17	$\pm$ 0.33	4
Lee[a]	–0.35 $\pm$ 0.20	$\pm$ 0.40	7
Teunissen *et al* (2011)	0.01 $\pm$ 0.32	$\pm$ 0.63	10
Brauer *et al* (1997)	–0.03 $\pm$ 0.42	$\pm$ 0.82	60
Al-Mukhaizeem *et al* (2004)	0.05 $\pm$ 0.22[b]	$\pm$ 0.43[b]	80
	Weighted Mean	$\pm$ 0.58	161

[a] In Byrne and Lim (2007).
[b] Weighted mean of three periods in table 1, Al-Mukhaizeem *et al* (2004).

to pulmonary arterial blood temperature where LoAs are ± 0.59 and ± 0.78 °C respectively (Lefrant *et al* 2003).

We also computed the RMSE for each individual volunteer and computed the mean RMSE $\pm$ SD for each condition. A single factor (study condition) analysis of variance was used to test for differences in KF model performance (RMSE, bias, and LoA) across conditions in study A and across field studies E through I. To readily identify what factors were causing main effect differences we used the least significant difference post-hoc test. T-tests were used to examine differences in performance between laboratory baseline measures and dehydration, acclimation and clothing configurations studies B, C and D respectively. An overall RMSE was computed, weighted by each individual and study duration. Overall bias and LoA were computed from all data points. Grubbs (1969) outlier detection test was used to identify RMSE and bias measures that differed significantly from each study's group responses.

3. Results

3.1. Model development

Figure 1(A) shows the discrete probability distribution of ΔCT used for the time update model and (B) a scatter plot of all CT by HR points showing the mean HR $\pm$ SD for

CT binned by 0.1 °C intervals. The discrete probability distribution mean was found to be 0.001 ± 0.022 °C min^{-1}. The regression of previous CT with current was found to be $CT_t = 0.9984 \cdot CT_{t-1} + 0.0622$ with an $r^2 = 0.99$. With the mean of the discrete probability distribution close to zero, and the regression coefficient close to one, and because we expect it to be equally likely that CT will either increase or decrease we set our time update model to $a_1 = 1$, $a_0 = 0$ and $\gamma = 0.022$.

The optimal piecewise line segment points that provided the best RMSE (0.27 ± 0.10 °C) and largest number of points within ± 0.58 °C ($96.1 \pm 6.7\%$) are shown in figure 1(B). A quadratic fit to these points defines the observation mapping function as $b_0 = -7887.1$, $b_1 = 384.4286$, and $b_2 = -4.5714$. The mean SD for the binned HR = 18.88 ± 3.78 beats/min so σ is set to 18.88. To keep the KF model simple the positively (low CT) and negatively (high CT) skewed HR distributions at the extremes of CT are ignored. Keeping the assumption that HR is normally distributed across all CT has the effect of slightly under and over estimating the rate of rise of CT for low and high CTs respectively.

The extended KF model is described in equations (3)–(8) (below) and provided as a Matlab script in appendix B. The model is based on the linear algebra equations from Welch and Bishop (1995) which we have simplified for use with scalar variables. For our analysis we selected a starting CT_0 value from the data and set the initial variance (v_0) = 0. At each new one minute time point (t) the six equations were applied iteratively to provide a new estimate of CT_t and its associated variance (v_t) given a current observation of HR_t, the previous CT_{t-1} estimate, and previous variance (v_{t-1}), thus:

(1) Compute a CT preliminary estimate ($\hat{CT}_t$) based upon the previous CT estimate (CT_{t-1}) and the time update mapping function (a_1 and a_0).

$$\hat{CT}_t = a_1 \bullet CT_{t-1} + a_0 = 1 \bullet CT_{t-1} + 0$$
$$\hat{CT}_t = CT_{t-1} \tag{3}$$

(2) Compute a preliminary estimate of the variance of the CT estimate ($\hat{v}_t$) based upon the previous CT variance (v_{t-1}) the time update mapping function (a_1) and variance (γ^2).

$$\hat{v}_t = a_1^2 \bullet v_{t-1} + \gamma^2 = 1 \bullet v_{t-1} + 0.022^2$$
$$\hat{v}_t = v_{t-1} + 0.000\,484 \tag{4}$$

(3) Compute the extended KF mapping function variance coefficient.

$$c_t = 2 \bullet b_2 \hat{CT}_t + b_1 = 2 \bullet -4.5714 \bullet \hat{CT}_t + 384.4286$$
$$c_t = -9.1428 \bullet \hat{CT}_t + 384.4286 \tag{5}$$

(4) Compute the Kalman gain (k_t) weighting factor based on the preliminary estimate of variance and using the extended KF variance coefficient.

$$k_t = \frac{\hat{v}_t c_t}{c_t^2 \hat{v}_t + \sigma^2} = \frac{\hat{v}_t c_t}{c_t^2 \hat{v}_t + 18.88^2}$$
$$k_t = \frac{\hat{v}_t c_t}{c_t^2 \hat{v}_t + 356.4544} \tag{6}$$

(5) Compute the final estimate of CT_t using the preliminary time update estimate, the error between the HR_t observation and the expected HR given the preliminary estimate of CT:

$$CT_t = \hat{CT}_t + k_t \left(HR_t - \left(b_2 \bullet \hat{CT}_t^2 + b_1 \bullet \hat{CT}_t + b_0 \right) \right)$$
$$CT_t = \hat{CT}_t + k_t \left(HR_t - \left(-4.5714 \bullet \hat{CT}_t^2 + 384.4286 \bullet \hat{CT}_t - 7887.1 \right) \right) \tag{7}$$

(6) Compute the variance of the final CT estimate (v_t):

$$v_t = (1 - k_t c_t) \hat{v}_t \tag{8}$$
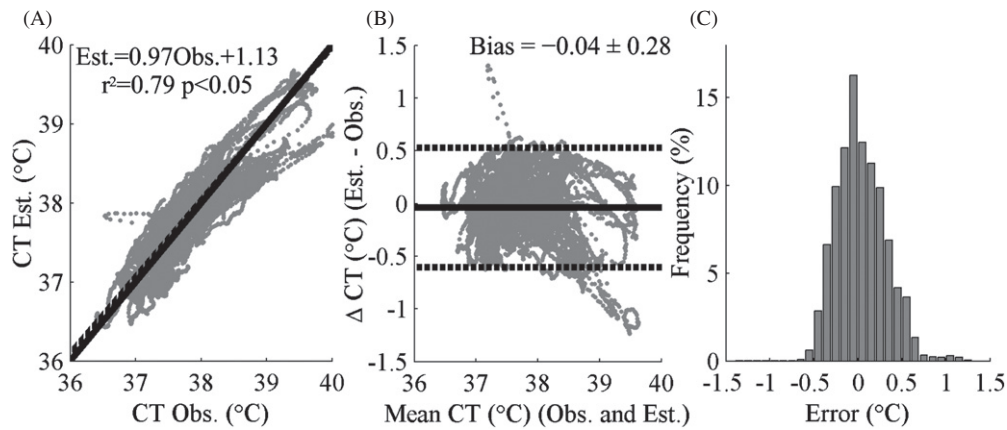
Figure 2. (A) Scatter plot of observed (Obs.) CT versus estimated (Est.) CT for the development data, showing the line of identity (solid) and least squares regression line (dashed). (B) Bland–Altman plot showing bias (solid) and ± 1.96 SD (dashed) for the development data. (C) Normalized histogram of model error for all training data.

Table 3. Use of the extended KF equations on a portion of the observed HR data.

HR	t	$\hat{C}T_t$ (equation (3))	$\hat{v}_t$ (equation (4))	c_t (equation (5))	k_t (equation (6))	CT_t (equation (7))	v_t (equation (8))
	0					**37.94**	**0**
124	1	37.940 00	0.000 48	37.550 77	0.000 05	37.940 31	0.000 48
111	2	37.940 31	0.000 97	37.547 91	0.000 10	37.939 62	0.000 96
119	3	37.939 62	0.001 45	37.554 27	0.000 15	37.939 79	0.001 44
145	4	37.939 79	0.001 92	37.552 66	0.000 20	37.945 25	0.001 91

Bold font = observed or initialization data. Equations (3)–(8) are applied iteratively to compute CT_t and v_t.

Table 3 shows the iterative application of these equations to a series of HR observations given a starting $CT_0 = 37.94\,^\circ$C and a starting variance of $v_0 = 0$.

Figure 2 illustrates the performance of our learned model on the development data. Panel (A) shows a scatter plot of estimated CT by observed CT, the line of identity and a least squares linear regression fit to the development data. Panel (B) shows a Bland–Altman Plot of mean of observed and estimated CT versus estimated – observed CT. The bias $= -0.04 \pm 0.28\,^\circ$C with the LoA $= \pm 0.55\,^\circ$C. Panel (C) shows a normalized histogram of the model error.

3.2. Model validation

The KF model was validated against 150 individual test sessions with 83 different volunteers (>52 000 CT observations) and had an overall bias of $-0.03 \pm 0.32\,^\circ$C with the LoA $= \pm 0.63\,^\circ$C. The overall weighted mean RMSE was $0.30 \pm 0.13\,^\circ$C. Figure 3(A) shows a scatter plot of estimated CT by observed CT, the line of identity and a least squares linear regression fit to the validation data. Figure 3(B) shows a Bland–Altman Plot of mean of observed and estimated CT versus estimated–observed CT of the validation data. Figure 3(C) shows a normalized histogram of the model error for all the validation data.

Table 4 presents the mean RMSE, bias, and LoA for estimated versus observed CT for the laboratory and field validation studies. Appendix A shows individual Bland–Altman
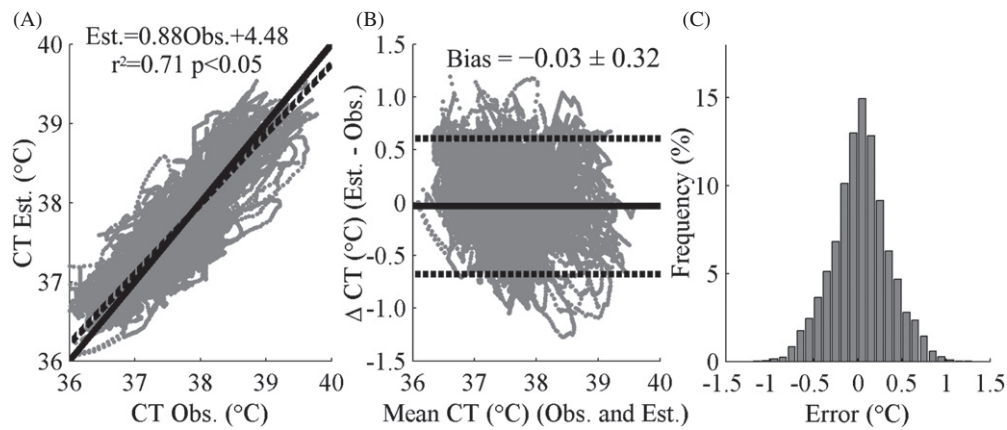
Figure 3. (A) Scatter plot of observed (Obs.) CT versus estimated (Est.) CT for the validation data, showing the line of identity (solid) and least squares regression line (dashed). (B) Bland–Altman plot showing bias (solid) and ± 1.96 SD (dashed) for validation data. (C) Normalized histogram of model error for all validation data.

Table 4. Mean RMSE, bias, and LoA for validation data.

Study	Condition	No. of min	n	RMSE	Bias	LoA
A.1. Environment	20 °C, 50% RH, 460 W	507	9	0.32 ± 0.16	−0.12 ± 0.33	**± 0.65**
A.2	27 °C, 40% RH, 350 W	461	11	0.25 ± 0.14	−0.09 ± 0.27	± 0.53
A.3	27 °C, 40% RH, 470 W	461	10	0.32 ± 0.13	0.07 ± 0.33	**± 0.65**
A.4	35 °C, 30% RH, 350 W	461	12	0.33 ± 0.18	**−0.25 ± 0.28[a]**	± 0.54
A.5	35 °C, 30% RH, 470 W	461	7	0.25 ± 0.12	0.07 ± 0.26	± 0.52
A.6	40 °C, 40% RH, 360W	461	7	0.29 ± 0.09	0.00 ± 0.30	**± 0.60**
B.1. Hydration	Hydrated	121	8	0.44 ± 0.19[a]	**0.31 ± 0.36**	**± 0.71[a]**
B.2 (33 °C, 50%)	Hypohydrated	121	8	0.26 ± 0.11	0.14 ± 0.24	± 0.48
C.1. Clothing	Shorts and t-shirt	111	6	0.21 ± 0.11	0.05 ± 0.23	± 0.45
C.2 (35 °C, 55%)	Chem. Bio. PPE	28	8	0.19 ± 0.16	−0.12 ± 0.21	± 0.40
D.1. Acclimation	Heat acclimated	59	7	0.28 ± 0.11	−0.13 ± 0.27	± 0.52
D.2 (45 °C, 20%)	Unacclimated	100	7	0.26 ± 0.19	−0.01 ± 0.31	**± 0.60**
E. US Army Rangers (24 °C, 85%)		140	11	0.29 ± 0.09	−0.06 ± 0.30	± 0.58
F. US Special Forces (SF) (11 °C, 91%)		1441	7	0.29 ± 0.07	0.06 ± 0.29	± 0.56
G. USMC Iraq (42 °C, 11%)		225	8	0.23 ± 0.08	−0.05 ± 0.24	± 0.48
H. USMC Afghanistan (20 °C, 20%)		586	8	0.32 ± 0.14	−0.07 ± 0.34	**± 0.66**
I.1. Austral. Sol. (MOPP II) (18 °C, 75%)		297	8	0.26 ± 0.07	0.03 ± 0.27	± 0.53
I.2. Austral. Sol. (MOPP IV) (18 °C, 72%)		244	8	0.42 ± 0.14[a]	**−0.28 ± 0.34[a]**	**± 0.67[a]**
Overall[b]				0.30 ± 0.13	−0.03 ± 0.32	**± 0.63**

Values are mean ± SD.

[a] Significant difference at $p < 0.05$. PPE = Personal protective equipment. Bolded results indicate Bias and LoA thresholds have been exceeded. USMC = US Marine Corps.

[b] Overall RMSE weighted by study duration and n. Overall bias and LoA computed from all data points.

(figure A1) and mean observed and estimated CT (figure A2) plots for all the studies to provide more detailed overview of the model performance.

At environmental conditions 35 °C, 30% RH, and an EE rate of 350 W (Study A.4) the bias exceeded our acceptability threshold and is significantly more negative than the study conditions A.3, A.5, and A.6 ($F = 2.77$, $P < 0.03$). The hydrated baseline condition (B.1)

Table 5. Individuals with RMSE and/or bias identified as outliers from two-tailed Grubbs test.

Individual (age, ht., wt., %fat)	Study	Outlying RMSE (°C)	Outlying bias (°C)
23, 1.70 m, 69 kg	A.2, A.3, A.4	0.60[a], 0.58, 0.74[b]	−0.59[a], −0.55[b], −0.72
*, 1.73 m, 72 kg, 9%	C.2	0.48[b]	−0.39
38, 1.86 m, 98 kg, 28%	D.1, D.2	0.38, 0.54	0.29[a], 0.50[b]
22, 1.85 m, 88 kg, 15%	H	0.60[b]	−0.55[a]

*Individual age not available.
[a] $p < 0.05$.
[b] Approaching significance.

Table 6. KF model performance for a variety of temperatures and energy expenditure rates with acclimated, hydrated volunteers, not encapsulated in personal protective equipment.

		Environmental temperature (°C)				
		9–13	18–20	24–27	33–35	40–45
Energy Expenditure Rate	Low (< = 375 W)	O	O↑	O	−	O↑
	Moderate (376–525 W)		O↑	O↑	O	
	High (526–675 W)		O	O	O	O
	Very high (>675 W)				+↑	

O = bias is < ± 0.13, ↑ LoA exceeds ±0.58 by less than 0.1 °C, ↑ LoA exceeds ± 0.58 by more than 0.1 °C, − underestimation of CT, + overestimation of CT, gray area = no data.

exceeded both the bias and LoA criteria, and is significantly different from the dehydrated condition (B.2) on the measures of RMSE and LoA ($t = 2.21$, $P < 0.05$; and $t = 3.05$, $P < 0.01$ respectively). For the field studies mild conditions (18 °C) with high EE rate (∼685 W) and encapsulation in PPE (I.2) has significantly greater RMSE, and negative bias ($F = 4.24$, $P < 0.004$, $F = 3.78$ $P < 0.007$ respectively) than the other filed studies (E, F, G, H, and I.1); and significantly greater LoA than studies G and I.1 ($F = 2.68$ $P < 0.03$, respectively).

Table 5 presents four individuals identified as outliers using the Grubbs criterion test from 7 of the studies. No individual characteristic stands out as a factor in determining the outliers.

Table 6, summarizes the performance of the KF model across a range of temperatures and TEE rates with clothing configurations from shorts and t-shirts to partial encapsulation.

4. Discussion

The KF model has an overall bias of only $−0.03 \pm 0.32$ °C and LoA of ± 0.63 °C, indicating that 95% of all KF model estimates fell within this range of the observed CT. The KF model has a similar LoA to those found when comparing rectal and esophageal measurements of CT (± 0.58 °C see table 2) and is within the LoA for rectal and esophageal measures found by Teunissen *et al* (2011) and Brauer *et al* (1997) of ± 0.63 and 0.82 °C respectively.

Using the aggregated results of the various studies, with clothing configurations from shorts and t-shirts to partial chemical biological encapsulation; it is possible to examine the performance of the KF model across a large temperature range for several different rates of energy expenditure. Table 6, summarizes the performance of the KF model in terms of our comparison criteria. For most temperatures and work rates the KF model provides CT estimates with a bias and LoA that are similar to those found in comparisons of rectal versus esophageal temperatures. However, at temperatures of 33 to 35 °C there are two exceptions.

First, at low work rates the KF model significantly underestimates CT, and second at very high work rates the model significantly overestimates CT. These errors can be tolerated in the context of using the KF model for maintaining the safety of individuals. At low work rates the underestimation poses limited risk for missing individuals under thermal strain. In fact, the CT observations for this 8 h series of work rest cycles (study A.4) never exceeded 38.5 °C. Conversely, at very high work rates the KF model tends to err on the side of false positives rather than missing thermally stressed individuals.

Analysis of the laboratory studies also demonstrates that the model provides CT estimates with small bias and LoAs within or close to our comparison threshold when volunteers are dehydrated, encapsulated in chemical biological PPE, or in an unacclimated state.

Only one set of conditions proved difficult for using the model in a safety assessment context. The KF model performs significantly less-well estimating CT of individuals engaged in very strenuous activity, in cooler temperatures while encapsulated in chemical biological PPE (MOPP IV) (study I.2). This particular study examined volunteers on a 5 km road march conducted at a 15 min mil^{-1} pace in full chemical biological protective garments. Here the model clearly does not account for the full rise in CT seen during the road march (see figure A2, panel I.2 in appendix A) and hence has a large negative CT bias. Although the work rate and clothing vapor occlusiveness are similar to that in the clothing laboratory study (study C), the ambient temperature was cooler (18 °C versus 35 °C). Under these conditions it appears that the thermoregulatory response of the volunteers was to widen the CT-to-skin temperature gradient, by allowing CT to rise, rather than increase skin blood flow (see Sawka and Young 2006). Thus, the observed rise in CT was greater than the rise our model would estimate from HR. Under the warmer conditions of the clothing laboratory study (study C) our model performed adequately.

When compared to the other recent approaches at providing non-invasive estimates of CT the KF model performs well. The KF model estimates of CT have a similar bias when compared to the heat flux sensor proposed by Gunga *et al* (2008). However, LoAs for the heat flux sensor in environmental conditions of 25 and 40 °C were much higher ($\pm$ 0.71 °C and $\pm$ 0.74 °C respectively) than all our conditions except study B.1. Similarly, when the KF model is directly compared to a real-time implementation of a physics based thermo-regulatory model (Yokota *et al* 2008) across study conditions A.1–A.5 (bias $\pm$ LoA: -0.24 ± 0.67; -0.25 ± 0.66; -0.08 ± 0.77; -0.23 ± 0.65; 0.10 ± 1.09 °C, respectively) (DeGroot *et al* 2008 personal communication) the KF model performs better with a bias closer to zero in four out of the five conditions and has LoAs less than those provided by the thermoregulatory model.

As with any method there is a distribution in performance (see figure 3) where the model will more accurately estimate CT in some individuals than others. The overall and individual study Bland–Altman charts in combination with the outlier analysis show that the KF model predicts CT very well except for a small number of individuals. The outlier analysis (see table 5) identified four individuals where the model did not perform as well as the group. This same number is predicted by the LoA methodology (5% of 83) where 5% of the population would be expected to fall outside of the ± 0.63 °C LoA bounds. For the outlier from study A.2, A.3, and A.4 the error is systematically negative. Similarly, the individual identified as an outlier in D.1 and D.2 has a systematic positive bias in CT estimations. With these outliers performance of the KF model appears to be individual specific rather than condition specific, and so systematic biases for individuals appear to be correctable. The model parameters do allow for individualization by both age and fitness and future research will examine how individualizing the model for these parameters can improve CT estimation. As most volunteers

in our data were drawn from relatively young and fit male military populations, further work will also be necessary to expand the model's generalizability to females, and older and less fit populations.

Other factors can affect HR, and thus our CT estimation, such as diet, caffeine, sleep, and psychological stress. The effects of diet and caffeine on HR will likely be outweighed by activity and thermoregulation. There were no controls for these factors in the field studies. But, since it is likely that many volunteers were taking caffeinated products, the performance of the model includes the potential influence of caffeine. During sleep HR is reduced to low levels, but the KF model's estimation of CT appears appropriate in these situations given the sleep data at the mid-point of Study F. However, the impact of elevated HR from sustained psychological stress on the KF model's performance is unknown. While, studies G and H contain active military patrols, further research is necessary to examine the effects of sustained stress on the model's performance.

Although in some conditions the KF model provides LoAs that exceed our comparison threshold, it is important to highlight both the simplicity of the present KF model and that the LoA calculations include all data with HR transients from the start and end of exercise. These transient periods are included to demonstrate that the model can track CT during dynamic real-world periods of work and rest. Teunissen *et al* 2011 examined the LoA between rectal and esophageal CT across periods of rest, exercise and recovery and found a similar LoA of ± 0.63 °C. Our KF model is simple using only one variable, HR, to estimate CT, with no adjustment for height, weight, body composition, fitness, or age.

Finally, while the model is not a replacement for the direct measurement of CT the findings suggest the KF model is accurate, precise and practical enough to provide an indication of thermal work strain for use in the work place. Using an estimate of CT in conjunction with HR to obtain a physiological strain index value would provide a simple mechanism for alerting workers to possible excessive thermal work strain.

5. Conclusion

Core body temperature can be estimated by a Kalman filter model using a single parameter, heart rate, to within similar bias and limits of agreement seen when comparing rectal and esophageal measurements of core temperature. The model was validated against a series of laboratory and field studies with 83 volunteers and 150 experimental runs across a range of environmental temperatures from 9 to 45 °C and work rates. We demonstrate that the model performs similarly in different environments, in the presence of dehydration, with limited or complete clothing occlusion, and whether volunteers are heat acclimated or not. While this technique is not a replacement for direct core temperature measurement, it offers a simple and promising new method for estimating individual core temperature and is accurate and practical enough to provide a means of real-time heat illness risk assessment.
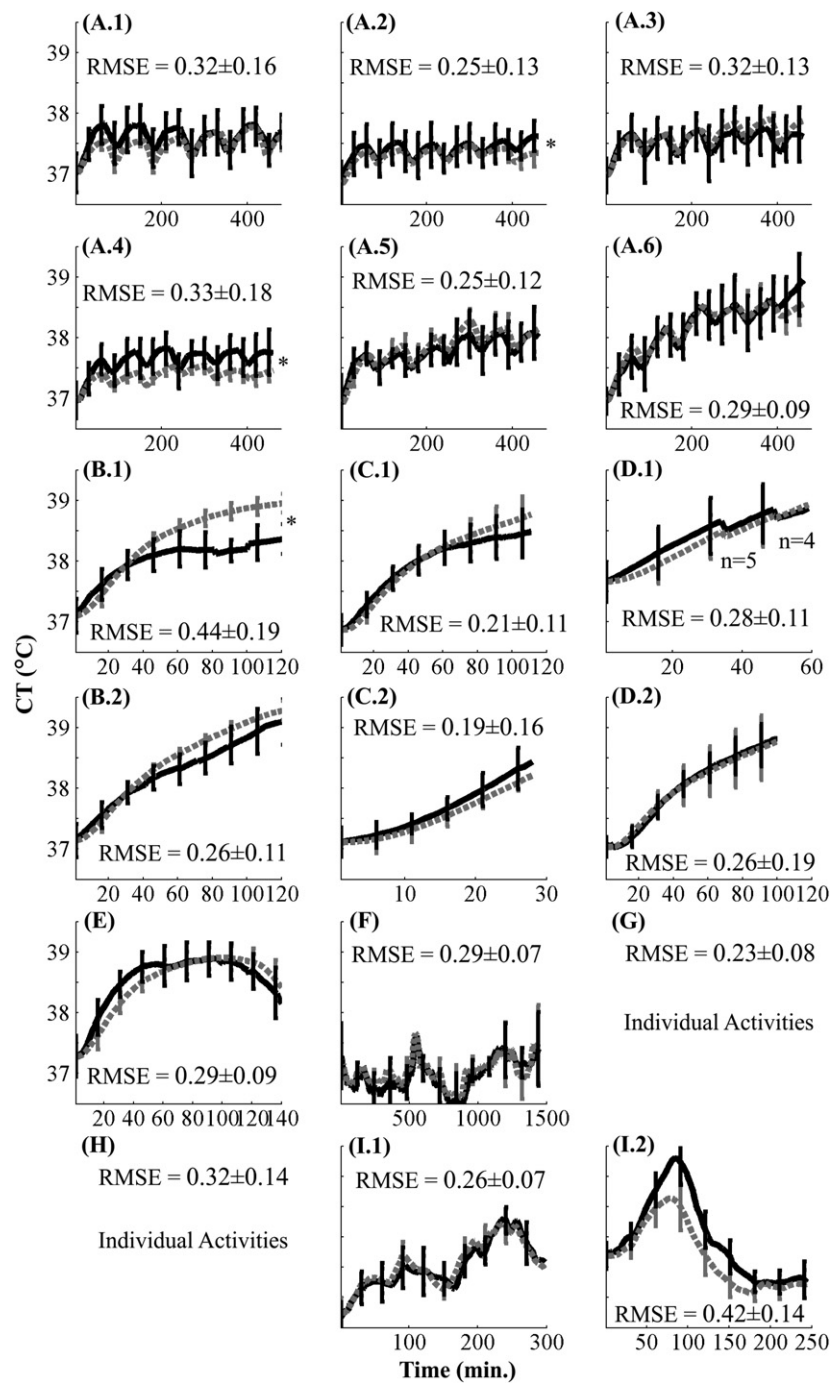
Acknowledgments

Appendix A. Bland–Altman plots and mean CT plots for individual studies
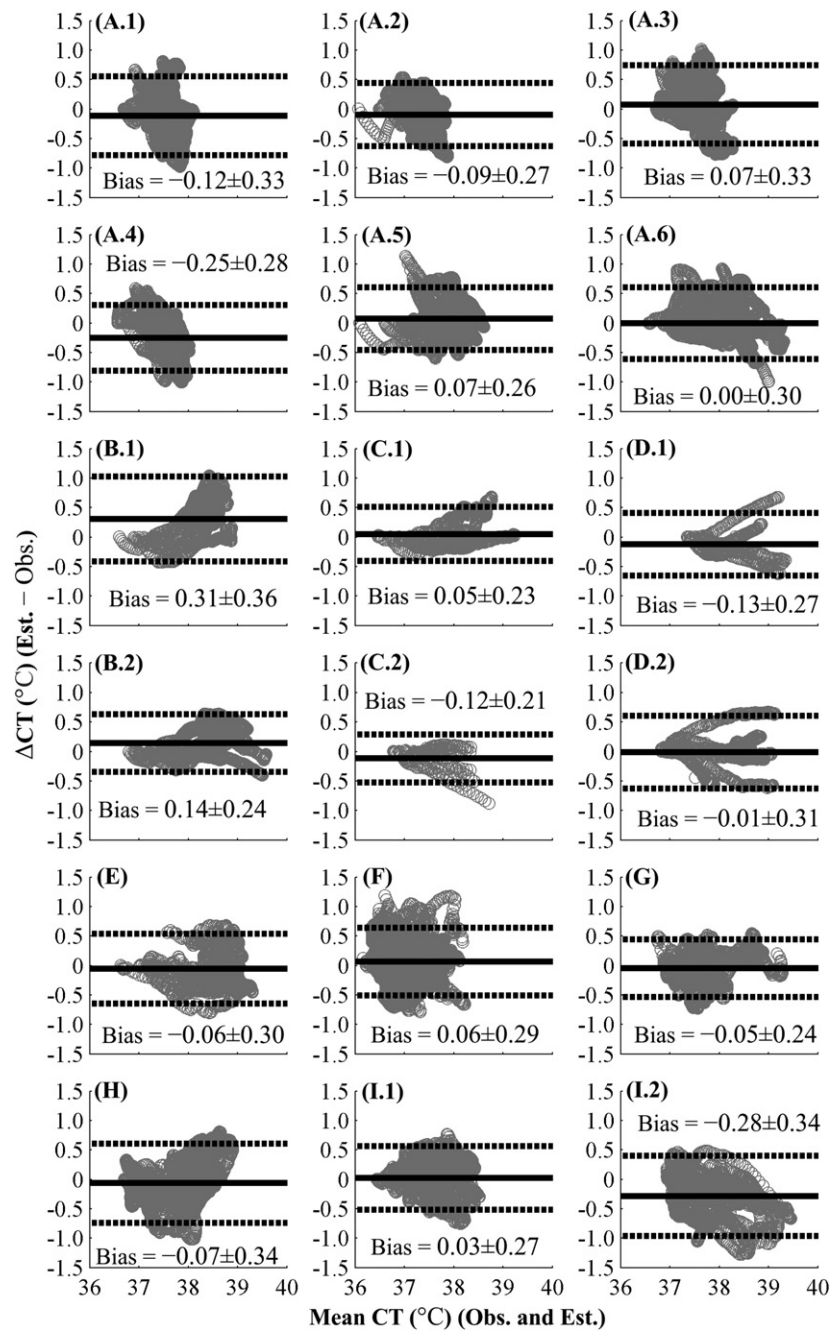
Figure A1. Bland–Altman plot showing bias (solid) and ±1.96 SD (dashed).

Figure A2. Mean observed CT (solid—black) and mean estimated CT (dashed—gray) with ± 1 SD. Means for studies G and H are not shown as they are a combination of several activities over the study period. * = end points significantly different $p < 0.05$.

Appendix B. Matlab function

```
function CT = KFModel(HR,CTstart)
%Inputs:
  %HR = A vector of minute to minute HR values.
  %CTstart = Core Body Temperature at time 0.
%Outputs:
  %CT = A vector of minute to minute CT estimates
%Extended Kalman Filter Parameters
  a = 1; gamma = 0.022^2;
  b_0 = -7887.1; b_1 = 384.4286; b_2 = -4.5714; sigma = 18.88^2;
%Initialize Kalman filter
  x = CTstart; v = 0;%v = 0 assumes confidence with start value.
%Iterate through HR time sequence
  for time = 1:length(HR)
  %Time Update Phase
   x_pred = a*x;                                    %Equation 3
   v_pred = (a^2)*v+gamma;                          %Equation 4
  %Observation Update Phase
   z = HR(time);
   c_vc = 2.*b_2.*x_pred+b_1;                       %Equation 5
   k = (v_pred.*c_vc)./((c_vc.^2).*v_pred+sigma);   %Equation 6
   x = x_pred+k.*(z-(b_2.*(x_pred.^2)+b_1.*x_pred+b_0)); %Equation 7
   v = (1-k.*c_vc).*v_pred;                         %Equation 8
   CT(time) = x;
end
```

References

Al-Mukhaizeem F, Allen U, Komar L, Naser B, Roy L, Stephens D, Read S, Kim C and Schuh S 2004 Comparison of temporal artery, rectal and esophageal core temperature in children: results of a pilot study *Pediatr. Child Health* **9** 461–5

Bland J M and Altman D G 1986 Statistical methods for assessing agreement between two methods of clinical measurements *Lancet* **1** 307–10

Braur A, Weyland W, Fritz U, Schuhmann M U, Schmidt J H and Braun U 1997 Determination of core body temperature. A comparison of esophageal, bladder, and rectal temperature during postoperative rewarming *Anaesthesist* **46** 683–8

Buller M J, Castellani J, Roberts W S, Hoyt R W and Jenkins O C 2011a Human thermoregulatory system state estimation using non-invasive physiological sensors *Proc. IEEE Int. Conf. on Engineering in Medicine and Biology Society* pp 3290–3

Buller M J, Tharion W J, Hoyt R W and Jenkins O C 2010 Estimation of human internal temperature from wearable physiological sensors *22nd Conf. on Innovative Applications of Artificial Intelligence (IAAI)* pp 1763–8

Buller M J, Welles A P, Stower J, Desantis C, Margolis L, Karis A J, Economos D, Hoyt R W and Richter M W 2011b Thermal work strain during Marine rifle squad operations in Afghanistan (March 2010) *USARIEM Technical Report* T11-02 (AD A501301)

Buller M J *et al* 2008 Thermal-work strain during Marine rifle squad operations in Iraq (summer 2008) *USARIEM Technical Report* T09-01 (AD B345494)

Byrne C and Lim C L 2007 The ingestible telemetric body core temperature sensor: a review of validity and exercise applications *Br. J. Sports Med.* **41** 126–33

Cheuvront S N, Montain S J, Goodman D A, Blanchard L and Sawka M N 2007 Evaluation of the limits to accurate sweat loss prediction during prolonged exercise *Eur. J. Appl. Physiol.* **101** 215–24

Consolazio C F, Johnson R E and Pecora L J 1963 Physiological variability in young men *Physiological Measurements of Metabolic Functions* ed C F Consolazio, R E Johnson and L J Pecora (New York: McGraw-Hill) pp 453–80

DeGroot D W, Goodman D A, Montain S J and Cheuvront S N 2008 Validation of the ICDA model for predicting body core temperature *Med. Sci. Sports Exerc.* **40** S367 Abstract

Fiala D, Lomas K J and Stohrer M 2001 Computer prediction of human thermoregulatory and temperature responses to a wide range of environmental conditions *Int. J. Biometeorol.* **45** 143–59

Fick A 1855 On liquid diffusion *Phil. Mag. J. Sci.* **10** 30–39

Fox R H, Solman A J, Isaacs R, Fry A J and MacDonald I C 1973 A new method for monitoring deep body temperature from the skin surface *Clin. Sci.* **44** 81–86

Frank A, Belokopytov M, Shapiro Y and Epstein Y 2001 The cumulative heat strain index—a novel approach to assess the physiological strain induced by exercise heat stress *Eur. J. Appl. Physiol.* **84** 527–32

Grubbs F E 1969 Procedures for detecting outlying observations in samples *Technometrics* **11** 1–21

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. Therm. Biol.* **33** 297–307

Gunga H C *et al* 2009 The double sensor—a non-invasive device to continuously monitor core temperature in humans on earth and in space *Respir. Physiol. Neurobiol.* **169S** S63–8

Havenith G 2001 Individualized model of human thermoregulation for the simulation of heat stress response *J. Appl. Physiol.* **90** 1943–54

Kalman R E 1960 A new approach to linear filtering and prediction problems *J. Basic Eng.* **82** 35–45

Kenefick R W, Cheuvront S N, Ely B R, Palombo L J and Sawka M N 2011 DEET insect repellent: effects on thermoregulatory sweating and physiological strain *Eur. J. Appl. Physiol.* **111** 3061–8

Kraning K K and Gonzalez R R 1997 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.* **22** 331–42

Latzka W A, Sawka M N, Montain S J, Skrinar G S, Fielding R A, Matott R P and Pandolf K B 1997 Hyperhydration: thermoregulatory effects during compensable exercise-heat stress *J. Appl. Physiol.* **83** 860–6

Latzka W A, Sawka M N, Montain S J, Skrinar G S, Fielding R A, Matott R P and Pandolf K B 1998 Hyperhydration: tolerance and cardiovascular effects during uncompensible exercise-heat stress *J. Appl. Physiol.* **84** 1858–64

Lee J K, Nio A Q, Lim C L, Teo E Y and Byrne C 2010 Thermoregulation, pacing and fluid balance during mass participation distance running in a warm and humid environment *Eur. J. Appl. Physiol.* **109** 887–98

Lefrant J Y, Muller L, de La Coussaye J E, Benbabaali M, Lebris C, Zeitoun N, Mari C, Saissi G, Ripart J and Eledjam J J 2003 Temperature measurement in intensive care patients: comparison of urinary bladder, oesophageal, rectal, axillary, and inguinal methods versus pulmonary artery core method *Intensive Care Med.* **29** 414–8

Lim C L, Byrne C and Lee J K W 2008 Human thermoregulation and measurement of body temperature in exercise and clinical settings *Ann. Acad. Med. Singapore* **37** 347–53

Montain S J and Coyle E F 1992 Influence of graded dehydration on hyperthermia and cardiovascular drift during exercise *J. Appl. Physiol.* **73** 1340–50

Moran D S, Shitzer A and Pandolf K B 1998 A physiological strain index to evaluate heat stress *Am. J. Physiol. Regul. Integr. Comp. Physiol.* **275** R129–34

Sawka M N and Young A J 2006 Physiological systems and their responses to conditions of heat and cold *ACSM's Advanced Exercise Physiology* ed C M Tipton, M N Sawka, C A Tate and R L Terjung (New York: Williams & Wilkins) pp 535–63

Steck L N, Sparrow E M and Abraham J P 2011 Non-invasive measurement of the human core temperature *Int. J. Heat Mass Transfer* **54** 975–82

Teunissen L P J, Klewer J, de Haan A, de Konig J J and Daanen H A M 2011 Non-invasive continuous core temperature measurement by zero heat flux *Physiol. Meas.* **32** 559–70

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, NC

Wilkinson D M, Carter J M, Richmond V L, Blacker S D and Rayson M P 2008 The effect of cool water ingestion on gastrointestinal pill temperature *Med. Sci. Sports Exerc.* **40** 523–8

Yamakage M, Iwasaki S and Namiki A 2002 Evaluation of newly developed monitor of deep body temperature *J. Anesth.* **16** 354–7

Yokota M, Berglund L, Cheuvront S, Santee W, Latzka W, Montain S, Kolk M and Moran D 2008 Thermoregulatory model to predict physiological status from ambient environment and heart rate *Comput. Biol. Med.* **38** 1187–93